

# CANADIAN JOURNAL OF PSYCHOLOGY

VOLUME 14, No. 3

SEPTEMBER, 1960

Psychology—becoming and unbecoming: G. H. TURNER.....	153
Meaning associated with the phonetic structure of unfamiliar foreign words: GORDON A. McMURRAY .....	166
Incidental learning in a simple task: DAVID QUARTERMAIN & T. H. SCOTT .....	175
Tactual and visual interpolation: a cross-modal comparison: A. V. CHURCHILL .....	183
Children's understanding of number and related concepts: P. C. DODWELL .....	191
The effects of non-reinforcement on response strength as a function of number of previous reinforcements: RONALD K. PENNEY .....	204
Distribution variables in simple discrimination learning in rats: M. R. D'AMATO .....	216
Some data relating to the possibility of using a shorter form of the Hebb-Williams test: J. J. LAVERY & D. BÉLANGER .....	220
Book reviews .....	224

PUBLISHED FOR THE  
CANADIAN PSYCHOLOGICAL ASSOCIATION BY THE  
UNIVERSITY OF TORONTO PRESS

AUTHORIZED AS SECOND-CLASS MAIL, POST-OFFICE DEPARTMENT, OTTAWA

# CANADIAN JOURNAL OF PSYCHOLOGY

*Editor:* JULIAN BLACKBURN

*Assistant Editor:* P. H. R. JAMES

THE CANADIAN JOURNAL OF PSYCHOLOGY is published quarterly in March, June, September, and December. Annual subscription, \$6.00, single number, \$1.75.

*Subscriptions.* Orders and correspondence regarding subscriptions, change of address, and purchase of back numbers should be sent to:

THE SECRETARY-TREASURER, CANADIAN PSYCHOLOGICAL ASSOCIATION

Box 31, Postal Station D., Ottawa, Ontario

*Contributions.* Original manuscripts and correspondence on editorial matters should be sent to:

THE EDITOR, CANADIAN JOURNAL OF PSYCHOLOGY

Queen's University, Kingston, Ontario

*Information for contributors.* The *Journal* publishes experimental and theoretical articles in all recognized fields of psychology. Contributors need not be members of the CPA or residents of Canada. Manuscripts, either in English or in French, should be submitted in duplicate, they should be double-spaced throughout, and they should follow standard practice as regards tables, figures, etc. Authors are asked to adopt the current practice in American and British journals of citing references by date, for example, Smith (1956), instead of by number. "Immediate publication" (i.e. in the next issue to go to press) can be arranged for authors willing to pay the extra costs involved.

## CANADIAN PSYCHOLOGICAL ASSOCIATION, 1960-61

*President:* R. B. BROMILEY, Toronto; *Past President:* G. H. TURNER, London; *President Elect:* R. B. MALMO, Montreal; *Secretary-Treasurer:* F. R. WAKE, Ottawa; *Employment Bureau:* MISS MARJORIE FAHRIG, 68 Tower Road, Ottawa 5.

The Canadian Psychological Association also publishes *The Canadian Psychologist*, the business and professional quarterly of the Association. Editor: W. R. N. BLAIR, 462 Melbourne Ave., Ottawa 3, Ont. Annual subscription, \$3.50; single number, \$1.00. Correspondence regarding subscriptions should be sent to the Secretary-Treasurer of the Association. Contributions should be addressed to the Editor.

## PSYCHOLOGY—BECOMING AND UNBECOMING\*

G. H. TURNER

*University of Western Ontario*

THERE ARE TABOOS in psychology and one of them seems to be the consideration of ultimate aims or ends. Such delicate matters are normally shunned in the professional literature and in conversation among psychologists, and one does not have to go far to find reasons. The most obvious is the prevailing philosophical climate of logical positivism or scientific empiricism, with its futile disavowal of metaphysics. Its more deliberate devotees undoubtedly eschew final purposes on the high grounds of personal conviction; others seems to do so out of either indifference to, or irresolution on, these weighty issues. Still others may know where they stand but regard it as a highly personal matter, or may fail to speak out of respect for the taboos. Whatever the reasons I regard this conspiracy of silence as tragic, and consequently I am going to make the question of ends the first major concern of my talk this evening.

Since ultimate aims or values receive such scant attention, I can hardly be accused of labouring the obvious if I take a moment to argue that we are lost without them in our professional as well as our private lives. How long can an intelligent clinical psychologist function without wondering, in a profound way, what he is trying to do? Is he there to help remove undesirable symptoms, to reduce psychic pain, to improve efficiency, to restore a former level of function, to contribute to growth or independence, or what? Does he see mental health in positive terms and, if so, what is the desirable state at which he is aiming?

The problem of the industrial psychologist is no different. Is his function to assist in the attainment of his client's goals, to increase production and efficiency, to improve the administrative process, to share his own perspectives with the client, to help in achieving community-oriented goals, or what? Division XIV of the American Psychological Association brought this problem into sharp focus for me by suggesting that graduate students in industrial psychology be made aware of the "realities" of business and industrial life. If I interpreted this statement aright, it meant that to be a satisfactory consultant you must know the goals of your client and help them attain them—a proposition that is by no means as innocent as it sounds! The university teacher and administrator faces the same difficulties. Would he revise the programme of studies? To what end? Is the goal to impart the core curriculum, to develop the currently fashionable skills, to turn out scientists, or what?

\*Presidential Address to the Canadian Psychological Association, Kingston, Ontario, June, 3, 1960.

Let us touch a more sensitive nerve and bring the problem closer to home. How can we order our own lives and maintain health and stability if we have no overriding purposes or values, if we are busily treading the activity cage to nowhere? Without some hierarchy of values life must be chaotic and pointless. Without awareness of our values, how can we accept responsibility for them and avoid being slaves to whatever early cultural influences fell to our lot? If we once grant the desirability of self-awareness, of integrity and synergy of thought, feeling, and action, then professional, scientific, academic, and private problems begin to merge into one, namely: What are our ultimate objectives and how may these best be implemented in our private and professional lives?

But what do we concern ourselves with at present? To avoid becoming involved in disputes over philosophy, or, worse still, religion, we will openly admit the pursuit of only the most banal and patently penultimate or antepenultimate purposes such as increasing production, efficiency, and profit, or, even better, improving employer-employee relations, or effecting better personal adjustment, or facilitating the process of socialization, or, better still, raising standards, in the profession or training better scientists, or, best of all, adding to knowledge or advancing psychology as a science and as a profession. How perfectly respectable, how normal, but how unsatisfactory. They all leave unanswered the question, why bother? To stop short of God, Mammon, the state, the fully actualized man, or some equally final alternative, is to admit that we do not know what we are doing or that if we know, we will not say.

Going the whole distance and specifying our ultimate purpose would, of course, end in nothing but words if we were not simultaneously under the necessity of retracing our steps and establishing the hierarchies of lesser purposes that in effect constitute its definition. It is only from such a hierarchy of aims that any clear implications for behaviour can emerge and by which behaviour may be consistently evaluated. Thus the healthy person who is intellectually alive and growing is cognizant of his goals, ever mindful of the need to revise them and always modifying his behaviour, the better to achieve them.

To expect us to reach agreement about ends is quite unreasonable, although much of our disagreement is, I am sure, semantic rather than real, but to be unwilling to be explicit about them seems indefensible. What greater barrier to effective communication could one devise? What more subtle form of misrepresentation could one practise? One can with honour admit indecision or confusion. But what can be said of those who would deny or ignore an issue?

To accept responsibility for deciding, no matter how tentatively, and for stating, no matter how provisionally, the final grounds on which we



base our lives might not seem to constitute a formidable assignment since the answers are already implicit in our behaviour. By our every action we commit ourselves to some purpose or other. But this is surely a taller order than I am making out. What is harder than to discover and make explicit one's assumptions? More than that, the values we assume, the ends implicit in our behaviour, do not enjoy some sort of separate independent existence. On the contrary they are firmly embedded in a view of man, in some conception of the nature of the universe and of man's place in it. What is his nature, what does he need, where is he going, how can we account for him? It is how we answer these questions that determines our hierarchy of purposes.

Only our total concept of man can provide us with the master cognitive map into which all our little maps must fit if we would have a coherent view of life. But how poorly supplied we are with comprehensive models, and the bits and pieces that constitute the conventional wisdoms are characterized by unrevealed assumptions, inconsistencies, grave omissions, and questionable validity. If we are so uncertain of man's nature, How can we be so sure we are doing what is best for him as consultants, counsellors, clinicians, and teachers? How can we be so confident we know the best questions to ask of nature in pursuing our study of man? As we are well aware it is hard to find good answers to poor questions.

Do not think that this is none of our business, or that it concerns only the incorrigibly speculative personality theorists. Do we waive all claim to a part in policy setting? Are we content to implement the decisions or others? If this is not our business, then what is?

An objection that seems to appeal to many is that it is not of immediate concern to us as scientists, because psychologists do not at present know enough to provide a coherent picture of man. Their rejoinder is "Give us a hundred years or so and we may have something of significance to offer." The case for ignorance is easy to make out but the argument from it would be much more convincing if it were not for the fact that the self-same psychologist-scientist lives the real life of a real person at home, in the community, and among his professional colleagues, fighting for what he thinks is right and making his way in a real world, on the basis of very real assumptions about the nature of man. (Rigorous experimentalists have, by the way, been known to scorn the less certain findings of their colleagues who study children, while confidently raising their own on principles handed down with the family silver.)

I grant that our knowledge is limited and that our concept of man must be rounded out with borrowed or improvised propositions. But surely we must continually incorporate such knowledge as we have. It would be the height of irony for a psychologist to devote his life to the

task of building the empirical foundations of knowledge about man while living his life on postulates that take no account of the knowledge already available in his discipline.

Let us make no mistake about it, our job is to throw light on the nature of man. As practical and realistic men, we must and do, in the interim, commit ourselves completely to those views about man we think are most worth the gamble. As psychologists, we owe it to ourselves, and to others, to say what particular assumptions we are prepared to act upon, in order to clarify our own thinking, to give more consistent direction to our action and, with respect to clients, to let them know what they are paying for.

If we do this we shall find ourselves no longer declaring knowledge to be for its own sake and no longer regarding psychology as an end in itself—a way of thinking that has crept into our textbooks and into articles discussing, for example, the importance to psychology of recruitment and proper selection, or the danger to psychology of early professionalization, and so on. This is only a manner of speaking, but perhaps it is an unfortunate one. Unintentional reification seems to develop easily into unintentional deification. While there is still so much question of psychology's effectiveness as a means, we hardly dare promote it to an end.

Man may need a lot of things but one thing he certainly needs is an adequate view of himself, one that is consistent with and grounded in his fullest knowledge of himself and of his world. This is where psychology fits in. The world does not go round in order to accommodate psychology. The study of psychology derives such significance as it has from its power to contribute to our knowledge of and perspective on man.

It may be well at this point to remind ourselves that psychology is by no means the only source of knowledge about and perspective on man. How easy it is to bow perfunctorily to the natural sciences, to neglect the other social sciences, to ignore philosophy and theology, and to completely overlook the rich resources of literature, history, and the arts. But how unwise and how ungrateful. Other disciplines are not merely "fruitful sources of hypotheses" but the sources of most of our concepts and nearly all of our new ideas. And where do we get the postulates on which we gamble our lives?

The point cries out for greater elaboration than I have the competence or time to give it here. I will content myself with the guess that before long psychology and the humanities will be on much better speaking terms and that even in secular institutions psychologists themselves may be teaching philosophy to their students for the same reasons that they now teach statistics, namely, because there is an important job to do that cannot safely be farmed out to other departments.

In this connection I was interested to discover that mutual practical

problems had brought together three theologians, a psychologist (Paul Meehl), and a psychiatrist, to undertake the serious task of comparing their respective views on man and of presenting, in collaboration, the points of similarity and difference in their positions, along with, be it carefully noted, the practical implications for counselling and therapy. The results were published in 1958 as a symposium under the title *What, Then, is Man?* It struck me as a conspicuous success and as an encouraging sign that we may be growing up. If theologians and psychologists can communicate to good effect on such troublesome matters, it should not be too difficult for psychologists to communicate with other psychologists if they really set their minds to it.

It is well to remember, then, that psychology can not claim sole possession of man. Moreover we should be encouraged by the thought that there are other shoulders to the wheel as well as ours. A great deal may depend on us but, fortunately, not everything.

We might also pause to reflect that marvellous as are man's cognitive functions they would be pale pathetic things without other modes of experience, and that we may too readily take for granted that science and knowledge and reason are all that matter in life. Just as we recognize in psychology the pitfalls in fragmenting the individual for purposes of study, we might also recognize the danger of fragmenting life for the purpose of living, as we do when we act as if *knowing* were everything. So when we return to a consideration of psychology, let us remember that it is psychology within science, in the context of human knowledge, within the context of human life. Not only are there other toilers in the vineyard but there are other things to be done besides studying the grapes. Perhaps this dubious metaphor may serve to summarize all I have said so far and also, if stretched to the breaking point, to convey all that I now plan to add.

My next observation, then, is that we do not seem to be sufficiently respectful of man, if one can judge by the theories currently in vogue. Certainly we pay little attention to his finest attributes and greatest achievement. And since personality theory, broadly conceived, provides a framework for living, for the selective organization of experience and for systematic scientific observation, I agree with Bronfenbrenner and others like him who feel that maturity of psychology as a science is peculiarly dependent upon the work we do in this area over the years ahead.

As far as current theorizing goes, I trace much of my dissatisfaction to certain of the presuppositions on which it seems to rest. To begin with, most of us seem to be epistemological empiricists determined to trace all knowledge to sense perception, as if, in some inevitable fashion, the data of experience combine themselves into coherent principles or understandings or generalizations. Speculation unrelated to and unrestrained

by the facts of observation has no place in modern psychology, but we must not let our enthusiasm for objectivity blind us to the fact that it is the active observer who apprehends the sense data and provides the frames of reference without which the facts would be meaningless. The implication I object to is that man brings nothing to a situation that he has not acquired through previous observations. This is just one of the ways in which we sell the psyche short, and also one of the ways in which we dichotomize our thinking and fail to appreciate the possibility of some intermediate position, such as, in this case, a combination of empiricism and rationalism.

To push this idea a little further we might well, for example, stop pretending that scientific insights are arrived at by pursuing a logical path of reasoning following a series of observations. A scientist's theories arise neither spontaneously from data nor from an application of logical modes of thought. His theoretical contributions rest on no such rational accomplishments, accessible to training, but upon the irrational and still mysterious resources of his own psyche. A scientific training in verification procedures (that is, in the testing of hypotheses), is no guarantee of the release or development of scientific creativity.

Combined with a thorough-going empiricism, we generally find a materialist monism, the idea that all psychic events are, in principle, reducible to physical events. For me, this is carrying the law of parsimony too far. It asks us to accept that the very psychic processes by which the world is apprehended are to be reduced to the objects held in apprehension. It is at very least a high-handed way to treat consciousness and is a further instance of the tendency to downgrade the human psyche. Dualism has its problems, but are they any more formidable than those inherent in monism? Dualism may be devilishly inconvenient, but it strikes me as being much more realistic. But more of this in a moment.

Our theories are also, for the most part, mechanistic. And after Ketchum's brilliantly persuasive article in the *Canadian Psychologist* on the virtues of mechanism, properly conceived, it might seem foolish of me to take issue. But he made two mistakes. The first was to refer to my favourites, the teleological self-realization psychologists, as a tribe; and I did not think he intended it as a compliment. From that moment, I knew that mechanism could not be permitted to go unchallenged. Certainly there is mystery enough in mechanism for anyone and certainly mechanism is essential as a methodological assumption, but why as a metaphysical one? The second mistake was, I feel, to assume that a biological vitalism is the only alternative to mechanism. Vitalism may or may not yet be dead, but emergentism and existentialism still have to be conquered and what guarantee there will not be others?

Since when did nature agree to conform to our simple requirements,

to conveniently fit our categories, to fill our concepts with nothing left over, to hew to this line, or to that? If she made no such contract, look for trouble. Think of what has happened in the natural sciences. On the one hand, we have, or so I am told, a three-billion-year-old universe with 100 million galaxies, some of them 180 million light years away and receding from us at the rate of 25,000 miles a second. On the other hand, when we get to what is presumably the essence of everything, the nucleus of the atom, we find it ungratefully composed of at least 21 different particles. Particles may, of course, behave like a wave and a wave of energy may act like a particle, but that is just to keep us on our toes. The nuclei containing all these particles are, at the largest, estimated to be about two-thousandths of a billionth of a millimetre, the small ones about 15 ten-thousandths of a billionth.

So much for space. Time too is a little crowded. If I am properly informed, a mu meson decays into an electron and two neutrinos in approximately two-millionths of a second, but this is a leisurely performance compared to that of the pi meson which zestfully decays into two gamma rays in a hundred-millionths of a second, only to be hopelessly outdone by the greatest speedster of them all, so far, the V particles which decay into a pair of positive and negative particles in ten-billionths of a second.

Energy too is not to be outdone. The nuclear force binding the particles in the nucleus of an atom together is sufficient to overcome the tremendous repulsiveness the particles have for one another, a repulsiveness so great that, as someone has estimated, two one-gram bundles of protons 8,000 miles apart would exert a force on one another equal to 28 tons.

The most wondrous thing of all is that this whole impressive panorama disappears into the nothingness of a space-time continuum leaving only mathematical abstractions residing in the mind of man. What a fate for materialist monism and for scientific empiricism! There is cool comfort here for those who would squeeze all of nature into their neat little categories and final philosophies. How humiliating if man should turn out to be so much simpler than one miserable little nucleus of an atom almost too small to think about! Do we really expect to be able to give an adequate account of man in one chapter of someone's future textbook on personality theory? So there's mystery enough in mechanism (to repeat Ketchum's quote from Hebb), but there will be enough left over to challenge the wisdom of ever attempting to rule out all postulates but our own.

Let us take another illustration, the perennially troublesome opposition of determinism and free will. Christians and all others who like to feel that their sense of freedom and responsibility is not an illusion are upset

by the claims of thorough-going scientific determinists; and social scientists, at least when they are functioning as social scientists, seem to panic at the suggestion, implicit in the doctrine of free will, that they have to add to their list of variables one that is belligerently, and in principle, uncontrollable. They are inclined to feel that if this sort of thing is going to be allowed to go on they might just as well pack up and go home—a decision for which they would, curiously enough, be prepared to take full credit, and responsibility. There is obviously some difficulty here that needs to be resolved. But the issue is important for other reasons. One cannot give a very coherent or comprehensive account of man without committing oneself on this matter and, in addition, the view held by the applied psychologist or other practitioner dealing with human problems, influences his attitude, understanding, treatment, or advice.

The scientific determinists would never have been quite so alarmed had they been fully aware of the views of the sophisticated advocates of free will who have never accorded man unlimited psychic freedom under all circumstances, but rather limited, if very important, freedom under certain optimal conditions; and all this within the framework of a very necessary determinism. Thus free will is not squarely opposed to determinism but is superimposed upon it, and the real adversary of determinism is seen to be indeterminacy. These ideas have been elaborated, with much more philosophical and psychological sophistication than I could possibly muster, by Father Mailloux in an article in the *Canadian Journal of Psychology* in 1953 that well repays careful reading. Feigl has taken essentially the same position in the brief reference he makes to the problem in the *American Psychologist* in March, 1959. But the substitution of indeterminacy for free will as the principal adversary has hardly had the effect of quietening the determinist's nerves. For a while he could, if getting the worst of it, threaten to appeal to the higher courts of the natural sciences. Then he made the unsettling discovery that, of all people, the physicists, those paragons of scientific respectability, had admitted to full membership in the family the impudent, rebellious little principle of indeterminacy, not, of course, without some reluctance and only with respect to the behaviour of the minutest of particles, but nonetheless with full privileges. Scientific determinism is then in a double dilemma. It has to make some kind of peace with freedom and with indeterminacy.

Fortunately, psychologists are beginning to recognize this problem and to do something about it. Paul Meehl suggests that we have three choices. The first he calls *methodological determinism*. This is the attitude that you are looking for, and expecting to find, laws in the



domain of human behaviour. If you find ones that hold strictly, all well and good. If you have to settle for probabilistic laws, that will be fine, as they will still be very useful. Obviously one can have all this and free will too.

The second type he calls *empirical* determinism. Here we have the somewhat harder view that since the assumption of determinism has paid off so handsomely and so much regularity has already been found in behaviour, apparent exceptions are almost certainly due to incomplete information, and thus the assumption has been empirically vindicated and will, henceforth, be firmly held unless, of course, impressive counter-evidence should appear. Such people can at least sit down and talk about free will without becoming emotional.

The third and last category is that of *metaphysical* determinism. This is simply the proposition that all psychological events take place according to universal laws, adhered to as a metaphysical presupposition in no sense dependent upon empirical evidence. Those who hold this view are the people who must, if they are consistent, confess without pride or shame that they may be enjoying life and glad to be part of it but can take no credit or blame for what they do, or do not do, and who should be ready to apologize if, owing to habits of long standing, they speak as if *anyone should* do anything.

Each of us must resolve it in his own way. I should just like to say that since we all talk as if, and act as if, we had some modicum of freedom, and since life would seem to be, for many, a rather bad joke if we did not have a little freedom, I rather wish we would be honest and consistent enough to provide for a little when going beyond our findings and telling our little story about man. It would not change by one iota the level of confidence of a single statistical difference. It might on the other hand bring our words into a little closer harmony with our behaviour and our experience, as well as have a desirable effect on our approach to many practical problems.

Perhaps the most significant point worth making about empiricism, materialist monism, mechanism, determinism and a number of other postulates or presuppositions underlying our theorizing about man is that neither philosophy nor science can demonstrate their necessity. They represent preferences, not some sort of final truth, and if they have the effect of limiting our methods of study and of restricting our theorizing about man we may abandon them or modify them as we wish. The most we will have to contend with will be social pressure since psychologists as well as others put great pressure on one another to conform. So let us abandon the notion that the little we know is the full measure of man and write him up in terms that do some justice to

the full range of his intellectual, emotional, and moral virtuosity. The picture should be consistent with our scientific knowledge but not limited to it.

There are three other common dispositions in psychological theory that I deprecate which are not so clearly of methodological significance yet which create both theoretical and practical difficulties and also contribute to the general devaluation of man. One can be traced to Descartes, the others are of much more recent origin.

The first, and oldest, is our inveterate habit of subdividing our subject, man, into a number of separate processes, isolated for more convenient study. As Snygg has reminded us, we leave out, in our study of process, the unifying principle, the person in whom the process resides, and then wonder why we cannot get the processes to fit together into a recognizable picture of a human being. For reference, see any introductory textbook in psychology. For this, and much more besides, read Father Malone's masterpiece in the April number of the *Canadian Psychologist*.

Our tendency to fragment our subject-matter is nowhere better illustrated than in the unfortunate splitting of psychological scientists and theory builders into three camps, the experimental-physiological enthusiasts, the devotees of social psychology, and the followers of personality dynamics. This division has led to a certain amount of misunderstanding, rivalry, and petty bickering. Certainly each has been inclined to look down his nose at the others (with predictable consequences) and so far attempts to resolve differences have not been very successful, possibly because the theoretical basis for union has been conspicuously lacking—at least until our little-known European friends, the existential psychotherapists, came along. They regard the individual as existing simultaneously and inseparably, in the biological, social, and self aspects of his world, and rather than pick any favourites they insist that none of the three be neglected. As all readers of Rollo May's *Existence* will know, their terms are, respectively, Umwelt, Mitwelt, and Eigenwelt and they mean essentially what Nuttin means by psycho-physical, psychosocial, and spiritual. This would seem to be a rather radical suggestion and whether or not the high-flying Umwelters will associate themselves with what are mistakenly perceived to be the lesser breeds remains to be seen. Certainly it is a major part of my thesis that it is imperative that they do so or we may never be able to put Humpty Dumpty together.

Also prominent among the assumptions implicit in much psychological discourse, but of much more recent origin, is that of the relativity of values. This not only embraces the fact that there are obvious differences between societies in the notions of what is right and wrong, but the more



fundamental denial that there could be universal values, unless at some distant date a universal culture should emerge. If we are prepared to regard healthy tissue and a strong ego as better than damaged tissue and a weak ego, then the gap between fact and value has been bridged and it becomes obvious that the necessary assumption of universal lawfulness carries with it the assumption of universal values. We know that men in different cultures satisfy needs in vastly different ways. But if they are satisfying the *same fundamental needs*, and if satisfying a need is a value, then such values must be universal. If this were not so, one could make sense neither of psychology nor of humanistic ethics.

This is by no means a trivial point, for a relativity of values as the only alternative to Christian ethics suggests that for the non-believers social disapproval is the only factor to be considered in evaluating one's conduct. We cannot be unmindful of the social effects of the theories we propound and to encourage such a conclusion as this is, in my estimation, both irresponsible and stupid.

In keeping with the conception of values as relative is the acceptance of conscience as essentially Freudian, as an introjection of parental standards. One cannot, of course, deny the reality of Freudian conscience, but one can deny it exclusive rights to the territory, especially as it is by no means the only alternative to a Christian conscience. Rogers hints at, and Maslow boldly postulates, an intrinsic conscience, a hypothetical construct with as much empirical validity as a good many respectable psychological concepts. Philosophically I suppose this represents a preference for ethical intuitionism rather than ethical empiricism, a preference I share for much the same reasons that I accept the universality of basic human values. But this is another question that is being dealt with much more profoundly by the existential psychotherapists with their concept of ontological guilt, a culture-free form of guilt that arises, at least in part, from the forfeiting of one's potentialities. They are now providing powerful secular reinforcement for what has been, in psychology, a rather thinly defended but crucially important position, that of the universality of human values.

There are only two points I still wish to make and these are closely related. One is that we place almost every goal ahead of the growth of the individual, and the second is that we conveniently ignore ourselves as the knowers and doers. I am struck by, and most uncomfortable about, the tremendous emphasis we give to knowledge, experience, and skill in our training of and evaluation of psychologists. This is, I suppose, just another consequence of the stress on objectivity and of our pursuit of almost every goal except the growth of the individual person. If a psychologist can give a Rorschach or use the correct test of significance,

what *he is*, is apparently irrelevant. He is presumably interchangeable with any other M.A. or Ph.D. in psychology who knows how to do the same things. Find me the text in applied psychology that tells us what we ought to *be* instead of what to do. It appears that a psychologist is a person who does *this* whereas a banker is a person who does *that*. As our existentialist friends tell us, we have become functions, not persons.

In school we work on the arithmetic and take a chance on how the child develops. We would not dare suggest that we might focus on the facilitation of growth and take a chance on the arithmetic. We know what we are doing. We are supposedly putting first things first. If pushed, we prefer a little monster who can calculate to a constructive "little character" who cannot. So it is with the training of psychologists, and I would suggest that when we ignore the character of the person, we are making a very serious mistake. No perfection of technique can overcome the baleful influence of an unhealthy destructive individual. Certainly we need skill and knowledge, but the healthy constructive practitioner at least has his own integrity to keep him from undertakings beyond his competence. Character and personality should come first because of their intrinsic worth, but an equally good case can be made for their instrumental value, not, I grant, when manipulating objects, but certainly whenever interpersonal relations are to any substantial degree involved. It seems quite clear that applied psychology is in great measure the question of the impact, on those seeking help with practical problems, of the individual person with his particular style of life and perspective on life: and his perspective on life is, of course, intimately bound up with his concept of man and of self. Thus, the most important single variable in any practical interpersonal situation is the total personality and outlook of the consultant, therapist, counsellor, teacher, supervisor, or whomever. Knowledge and skill must be developed within the framework of this understanding.

Most of the scientists among us will, I am sure, violently reject any such notion. Not only do they have no time for such frivolities, but many can be trapped into arguing that personality disabilities of one sort or another probably account for their creativity. They deduce from the numerous biographies of unstable geniuses that a healthy mature personality would probably be their undoing. I should like to dissociate myself from all such nonsense. If we do not have the wit to value our own personalities and lives above success at a job, we might at least wonder how research undertaken by a mature person free to devote all his energies to the task, undisturbed by irrelevant considerations of status, rivalry, pecuniary advantage, or by distracting anxieties and hostilities, can be improved by the imposition of such handicaps. And should it subsequently be demonstrated that good mental health is negatively

correlated with scientific productivity then I would unhesitatingly recommend that we lower the scientific productivity rate.

My last point has to do with a particularly difficult feature of the life of a psychologist, namely, that he is, in a very real sense, part of his own subject-matter. Consequently, all he learns about man must, if he is to maintain intellectual and emotional integrity, be applied to his own life. In most other disciplines new theories or new understandings may present a substantial intellectual challenge, but that is merely part of the job and affects only a technical segment of the scholar's thoughts and actions. The psychologist must not only be constantly modifying his theoretical viewpoint but simultaneously be making appropriate adjustments in his whole outlook on and style of life. If he commits himself to a non-directive philosophy, he must change some of his behaviour. Likewise if he becomes a Skinnerian. It gives a certain urgency and seriousness to the consideration of new ideas and is undoubtedly wearing, but not to do so is to be either dissociated or fraudulent. All signs point to the necessity of focusing attention on our own personal intentions, convictions, and behaviour. For the psychologist more than anyone else, it is essential that he concentrate on his own growth as the only sound basis for rendering any help to others.

In conclusion, I am very optimistic concerning the future of psychology, not simply on the basis of past achievements but because of the tremendous potential still latent in a scientific approach to the understanding of man and his problems. It is my conviction, however, that psychology's potential, both as a science and as an art, will be more fully realized if we are clearer about the ultimate objectives to which we are prepared to commit ourselves, if we are more problem- rather than method-oriented in our research, and if we are less inclined to base the inferences we draw concerning the nature of man on the pre-suppositions underlying our research methodologies and on a range of information restricted to our own field. We must, in other words, break out of the methodological and philosophical traps in which we have already become ensnared. As to what some of these traps may be and what we might be doing about them, I have just offered a few suggestions, or more accurately, I have just offered my selection from among the many suggestions that have been made by others. My selection has, of course, been determined by my own orientation which might best be described, at the moment, as that of a Maslowian humanism under the first impact of existentialism. In any event, I have indicated in what direction I am headed and my plea is not that you should come with me, although I would much enjoy the company, but that you should let the rest of us know where you are going. That arrival can never be guaranteed is no excuse for failing to reveal the proposed destination.

## MEANING ASSOCIATED WITH THE PHONETIC STRUCTURE OF UNFAMILIAR FOREIGN WORDS<sup>1</sup>

GORDON A. McMURRAY

*University of Saskatchewan*

IN THE INTRODUCTION to his paper, Sapir (1929) speaks of referential and expressive symbolism in language. Referential symbolism means the learned, specific, arbitrary associations by which various words acquire their denotative meanings in the course of the development of a language. Expressive symbolism refers to a basic symbolic process, spread across diverse languages, whereby both expressive intonation and the phonetic structure of words may suggest meanings. In order to demonstrate this second sort of symbolism, Sapir (1929) constructed pairs of "non-linguistic" words matched for the initial and final consonant while the vowels between were varied along the scale a, ä, ε, e, i (e.g., mal-mil). His subjects strongly agreed in regarding the word with the vowel near the a end of the series as symbolizing a relatively larger reference.

Newman (1933) expanded this study to investigate whether such symbolic connotations were carried by the position of the tongue, the frequency of the acoustic formants, or the size of the resonant cavity. In addition to the large-small dimension used by Sapir, he also included bright-dark as one of the scales along which the constructed words were judged, and found again significant agreement among subjects in the placement of the words along both dimensions.

Bentley and Varon (1933) suggested that these findings were due to the subjects discovering in the words devised by Sapir and Newman the same dimensions of volume and brightness which are attributes of tones, that is, their results could have been obtained because of this property of sounds without such a psychological mechanism as natural or expressive symbolism ever actually operating in language.

Brown, Black, and Horowitz (1955) investigated the possibility that such symbolism actually occurs in natural languages. They selected 21 pairs of English antonyms from the Thorndike-Lorge word list according to two criteria: (1) that they should name sense experiences and (2) that they fall in the frequency range of 100 or over per million. These words were translated into their foreign equivalents in Czech, Chinese, and

<sup>1</sup>This study was completed while the author held a Senior Fellowship from the Canada Council. The author wishes to thank the Canada Council and Professor Paul Fraise of the Laboratoire de Psychologie Expérimentale et Comparée de la Sorbonne, who generously offered the facilities of his laboratory during the period of data analysis and report.

Hindi. The English antonyms were then paired with their foreign equivalents, and student subjects were asked to match the two pairs: for example, subjects knew that bright-dark translated into chamakdar-dhundhala, but since the order of the foreign words was random they would have to decide whether chamakdar was better translated by bright and dhundhala by dark, or vice versa. The results indicated that the subjects were correct in their matching (that is, agreed with the translations) to a degree which was significantly better than chance.

It appears in these foreign words studied that some characteristics of their phonetic structure suggest meanings which for some reason often agree with their denotative meanings. In Sapir's terms, the referential and expressive symbolism of these words have coincided significantly more than could be expected by chance. Further clues regarding the basis of this agreement in matching foreign words with their English equivalents, which has been found so regularly by several investigators, may be furnished by asking for judgments against several paired-opposites instead of only one. For it is very likely that these foreign words do not have connotations only along the dimensions which they happen to match denotatively, but may be placed along other dimensions as well. Since Osgood's semantic differential scales appear to be well adapted to studying connotations of this sort, the present study was undertaken to investigate more exhaustively the dimensions of meaning that could be carried by a small sample of foreign words by asking for judgments along a series of scales. The use of several scales also permits the analysis of interrelations between these judgments to discover the underlying dimensions of connotative meaning.

#### METHOD

##### *Selection of Words and Scales*

Fifteen pairs of antonyms, five pairs to represent each of the three languages, Czech, Chinese, and Hindi, were selected from those studied by Brown, Black, and Horowitz (1955). These words were chosen as those on which their Ss had obtained the highest percentage of successful translations. The mean percentage agreement on correct translations for these 15 pairs was 81.6 in their first sample ( $N = 86$ ) and 83.6 in their second smaller sample ( $N = 16$ ). The percentage agreement in every pair was significant ( $p < .01$ ) in at least one of their two samples and, in most cases, both.

Twelve descriptive scales, each one defined by a pair of polar adjectives, were selected from Osgood's (1957) lists. These were chosen to be relevant to this study and to adequately represent the main factors identified in factor analyses of semantic differential ratings (Osgood, 1957). They were not used as the usual seven-step scales, however, but as dichotomous scales. Both the words and scales so chosen are shown in Table I.

### *The Semantic Differential*

A semantic differential was then prepared by rotating each pair of words systematically against each scale in such a way that successive ratings of any one pair were separated by 14 other ratings. The foreign word-pairs were arranged in two columns numbered 1 and 2, followed immediately by the two opposites of the dichotomous semantic scales. For example, items 24 and 25 of the differential appeared as

	1	2				
24	bahut	ek	fast	---	slow	---
25	hrom	blesk	low	---	high	---

### *Subjects and Instructions*

The Ss were 76 students from an introductory psychology class at the University of Saskatchewan. They were told that the foreign words had opposite meanings, and were asked to match them with the ends of the semantic scales by placing the numbers 1 and 2 in the appropriate spaces. For example, in item 24, if 1 were placed after fast and 2 after slow it would indicate that S felt that *bahut* should be matched with fast and *ek* with slow. Preference for the opposite orientation would be indicated by placing 2 after fast and 1 after slow. The words were not pronounced, but only read silently by each S. Since 15 word-pairs were rotated against 12 scales, there were 180 such judgments which Ss were instructed to mark rapidly, judging each item as independently as possible.

### RESULTS

The results are shown in Table I. The figures in the body of the table give the percentage agreement on each judgment. Positive values indicate that the orientation agreed upon for the word-pairs and semantic scales was the same as that in which they are printed; negative values mean that the opposite orientation was favoured. For example, with *dlohý-kratý*, 88 per cent agreed that *dlohý* went with blunt and *kratý* with sharp; whereas the -85 per cent in the next entry shows that 85 per cent agreed that *dlohý* went with dark and *kratý* with bright. Since 66 per cent agreement is required for significance at the .01 level, it will be seen that 113 of the 180 judgments reached this criterion. Obviously, subjects showed remarkable agreement in deciding how the word-pairs should be aligned with the semantic scales.

A Friedman two-way analysis of variance by ranks was done with the data of Table I. The analysis showed significant differences both between word-pairs and scales.  $\chi^2$  between word-pairs was 35.1,  $p < .001$ ; and between scales was 36.9,  $p < .001$ . When the words were grouped according to language, the same test showed no significant difference between the percentage agreement means. This result agrees with Brown, Black, and Horowitz (1955) who also reported no clear differences between languages.

TABLE I  
PERCENTAGE AGREEMENT IN MATCHING CZECH, CHINESE, HINDI WORD-PAIRS WITH SEMANTIC SCALES\*

Word-pair	Percentage agreement in matching word-pair with semantic scale												Mean
	blunt sharp	bright dark	weak strong	active passive	short long	hard soft	beautiful ugly	large small	dissonant harmonious	fast slow	low high	heavy light	
<i>Czech</i>													
tupý-špičatý	88†	-78	68	-87	53	-67	-67	58	-67	-85	88	85	74.2
dlouhý-krátký	88	-85	72	-89	-64†	-83	64	62	-75	-89	86	69	77.2
rychlý-pomalý	-76	81	-58	80	54	88	50	64	63	86†	-82	-67	70.8
tvrdý-měkky	76	-71	-68	-68	-78	51†	-51	81	59	-64	66	72	67.1
hrom-blek	79	-83	-59	-78	-67	-59	-56	70	-61	-80	75	80	70.6
<i>Chinese</i>													
k'uai-màn	-83	80	72	67	-59	55	68	-53	54	78†	-73	-74	68.0
an-liang	57	-58†	64	63	96	-59	-68	-74	58	63	57	-54	64.2
kāng-jou	-61	61	-58	61	-58	74†	-62	79	67	70	-70	59	65.0
mei-ch'ou	-67	79	71	76	58	58	64†	-68	51	82	-81	-80	69.6
chung-ch'ing	84	-82	-70	-88	-58	-59	-83	91	74	-89	88	95†	80.1
<i>Hindi</i>													
nahe men-sanjida	72	-85	-64	-62	-58	-54	-71	75	67	-64	80	71	68.6
bahut-ek	88	-88	-62	-74	-80	-68	70	85	79	-75	75	88	77.7
dhundhala-chamakdar	83	-84†	-51	-70	-63	-64	-61	59	-62	-77	75	75	68.7
lamba-chhota	75	56	72	-63	-68†	-75	75	-63	-72	-62	59	51	65.9
gothil-tez	86†	-84	-72	-59	-84	-61	61	82	-53	-86	65	78	72.6
Mean	77.5	77.0	65.4	72.3	66.5	65.0	64.7	70.9	64.1	76.7	74.7	73.2	70.7

\*Positive values indicate that word-pairs were matched with scales in the order printed; negative values indicate the opposite orientation.

†Cases where the semantic scales are actual translations of the corresponding word-pair.



Several scales, marked in Table I by a dagger, were actual translations of the corresponding word-pair. Under the conditions of this experiment the mean agreement for these 12 translations was 74.7 per cent, lower than the values of 81.6 and 83.6 obtained by Brown, Black, and Horowitz (1955), but still far beyond the chance level. It is of interest to note though that, according to the median test, this value of 74.7 per cent was not significantly better than the over-all mean agreement of 70.7 per cent. It appears as if subjects can match the word-pairs with many dimensions equally as well as they can with the dimensions which happen to have the same denotative meanings.

It will be noted in Table I that the data have a possible range from +100, indicating complete agreement in aligning word-pair with semantic scale in the order they are printed, to +51 indicating a weak preference in the group for this alignment. When the opposite orientation is favoured, the possible range is from -100 to -51. A score of 50 per cent means that the group divided evenly, that is, there was no preference for one orientation over the other. This means that the scale jumps from +50 to -50 with no intermediate values. In order to remove this discontinuity each score in Table I was reduced by 50; for example, a score of -83 becomes -33 and a score of +80 becomes +30. This converts the data to a continuous scale ranging from +50, indicating maximum preference for orientation of word-pair and semantic scale in the order printed, through zero, indicating no preference, to -50, indicating maximum preference in the opposite direction.

The data of Table I, so converted, were then analysed by the D-method of factoring as described by Osgood (1957). According to this method, successive dimensions (factors) were extracted, using as the pivot at each step that scale which had the highest sum of squares across the 15 word-pairs. By this criterion, blunt-sharp, large-small, beautiful-ugly were selected as the first, second, and third pivots, respectively. The unrotated dimension co-ordinates on the first three dimensions are shown in Table II.

The first to appear was a very generalized, dominant dimension with at least six scales having co-ordinates indicating close relationship with the dimension. These scales, arranged in descending order of the size of the co-ordinates, were blunt-sharp (pivot), bright-dark, fast-slow, low-high, heavy-light, active-passive. Something in the foreign word-pairs emerged as a contrast between sharp-bright-fast-high-light-active on the one hand and blunt-dark-slow-low-heavy-passive on the other. This contrast was not only the dominant one along which these subjects were able to judge the strange words, but the percentage agreement on the six scales associated with this dimension was significantly higher than on



TABLE II

UNROTATED DIMENSION CO-ORDINATES OBTAINED BY D-METHOD OF FACTORING SCALES *versus* CONCEPTS

Scales	Unrotated co-ordinates with dimension		
	I	II	III
blunt-sharp	112.8	0.0	0.0
bright-dark	-105.1	-12.6	15.2
weak-strong	-14.3	-44.5	14.1
active-passive	-86.5	-4.0	5.4
short-long	-43.1	-40.4	-43.9
hard-soft	-56.0	32.4	-13.2
beautiful-ugly	-11.6	-13.1	63.2
large-small	54.1	73.5	0.0
dissonant-harmonious	-9.4	42.1	-13.3
fast-slow	-103.3	-2.0	2.8
low-high	96.0	-3.5	-19.7
heavy-light	92.2	19.7	-19.0

Pivots one, two, and three were blunt-sharp, large-small, beautiful-ugly, respectively.

the remaining six, according to the median test which gave  $\chi^2 = 24.2$ ,  $p < .001$ .

The second to appear was a less generalized, weaker dimension, but with three scales having co-ordinates indicating close relationship. These scales, again arranged in descending order of the sizes of their co-ordinates, were large-small (pivot), weak-strong, dissonant-harmonious. Although the first dimension was the single major one, it was not the only factor operating, the second one appearing as a contrast between one pole at small-weak-harmonious and another at large-strong-dissonant. The three scales associated with this second dimension were not judged with significantly better agreement than the remaining three scales (still unassigned), according to the median test which gave  $\chi^2 = .71$ ,  $p > .30$ .

The third pivot was the scale beautiful-ugly, but at this step the scales, with the exception of short-long, had co-ordinates which indicated low relationship with the pivotal dimension or had higher co-ordinates on other dimensions. Consequently, the remaining scales beautiful-ugly, short-long, hard-soft did not appear to form a coherent dimension in this study.

The analysis was not carried beyond the third step when it appeared that only the pivot scale would show a high co-ordinate with the new dimension which could thus be regarded as a specific factor.

Figure 1 shows the position of each of the words on the two main dimensions. The values of the co-ordinates were obtained by taking the

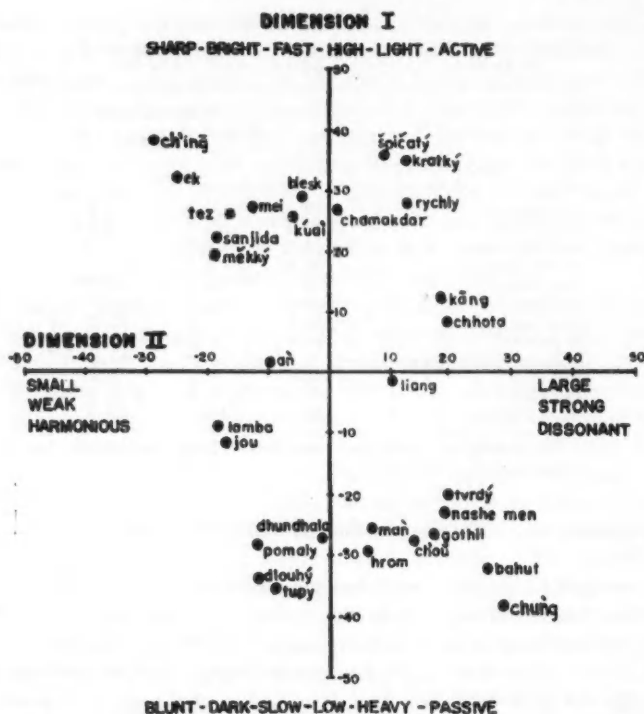


FIGURE 1. Relative placement of the word-pairs on Dimensions I and II.

mean of the agreement scores for each word-pair on the six scales most closely associated with Dimension I and the three scales most closely associated with Dimension II, respectively. The words are shown separately in the figure, but it should be noted that they were always judged in pairs and that the associated words are necessarily in symmetrically opposite positions.

#### DISCUSSION

The significant agreement of subjects in aligning foreign word-pairs with semantic scales appears as a clear result in this study. Presumably, something about the phonetic structure of these words, as suggested to English subjects reading them silently, possessed some common characteristics which formed the basis of agreement. These common characteristics could be said to form a dimension of judgment.

The results show that the dominant underlying dimension is the one defined by sharp-bright-fast-high-light-active, on the one hand, and blunt-dark-slow-low-heavy-passive, on the other. This contains many scales from Osgood's activity factor, and it may be that scales which ordinarily possess other connotations as well have swung into alignment with this dominant dimension. Subjects were also able to judge the words along the scales associated with this dimension with significantly higher agreement than on the others.

What is it about the words which places them on this dimension? Figure 1 may show the answer, for in the upper half of the figure (towards the positive pole of Dimension I) are found words in which the vowels are predominantly *ε*, *ai*, *ä*, *i*. In the lower half (towards the negative pole of Dimension I) are found words in which *a*, *ou*, *u*, *o* sounds predominate. It is also interesting to note that the word-pairs judged with highest agreement (*chuñg-ch'ing*, *bahut-ek*, *dlouhý-kratý*) are those in which the vowel contrast is marked. On the other hand, the word-pairs judged with least agreement are *añ-liang*, *kāng-jóu*, *lamba-chhota* in which the vowels are not as clearly separated. As in the earlier studies (Newman, 1933; Sapir, 1929), differences in vowels suggest different placement along various scales. The number of scales, however, was so restricted in these studies that the general nature of the underlying dimensions did not emerge.

Is it possible to proceed further with analysis of this sort? The results seem to support the taking of one more step. The dimension large-strong-dissonant versus small-weak-harmonious is not completely obscured by the dominant mode of variation. It is not judged with the degree of agreement of the first dimension, although still above chance. The scales associated with this dimension to some extent suggest Osgood's potency factor. It is also not easy to see what it is about the words that forms the basis for this differentiation. Inspection of the ++ quadrant of Figure 1 suggests that, possibly, words with hard consonant sounds like *kāng*, *kratý*, *špičatý*, *chamakdar* are moved towards the large-strong-dissonant pole. The vowels and consonants no doubt interact in a complex way. Also, the nature of the opposed word in each pair must be kept in mind when considering how the judgments are made.

The third dimension does not appear clearly enough to warrant further analysis. Different sampling of both words and scales might reveal other dimensions and give further indication of the essential features of the word structure that lead to different placement along these dimensions. In particular, the evaluative factor should be more adequately represented by additional scales. Although these refinements require further investigation, it seems clear now that differences in phonetic structure of words

do lead to different placement along dimensions of experience the general nature of which appears in this study.

### SUMMARY

Fifteen pairs of antonyms, five pairs to represent each of the three languages Czech, Chinese, and Hindi, were judged against each of 12 dichotomous semantic differential scales. Ss ( $N = 76$ ) showed very significant agreement in matching the foreign word-pairs with the scales. Analysis of the judgments showed one dominant dimension described as sharp-bright-fast-high-light-active *versus* blunt-dark-slow-low-heavy-passive and a second dimension large-strong-dissonant *versus* small-weak-harmonious. It was concluded that differences in the phonetic structure of unfamiliar words suggest meanings along these dimensions.

### REFERENCES

- BENTLEY, M., & VARON, EDITH J. An accessory study of phonetic symbolism. *Amer. J. Psychol.*, 1933, 45, 76-86.
- BROWN, R. W., BLACK, A. H., & HOROWITZ, A. E. Phonetic symbolism in natural languages. *J. abnorm. soc. Psychol.*, 1955, 50, 388-393.
- NEWMAN, S. S. Further experiments in phonetic symbolism. *Amer. J. Psychol.* 1933, 45, 53-75.
- OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana: Univer. Illinois Press, 1957.
- SAPIR, E. A study in phonetic symbolism. *J. exp. Psychol.*, 1929, 12, 225-239.

## INCIDENTAL LEARNING IN A SIMPLE TASK

DAVID QUARTERMAIN  
*Victoria University of Wellington,  
New Zealand*

AND

T. H. SCOTT<sup>1</sup>  
*University of Auckland,  
New Zealand*

INCIDENTAL LEARNING has been of theoretical interest mainly for its relevance to the problem of learning-motivation interaction. Early studies were designed to establish whether such learning occurred (Jenkins, 1933; Lepley, 1935; Postman & Senders, 1946). Later research has explored variables of which this learning is a function, usually in comparison with intentional learning. It has been shown that the differences between incidental and intentional learning are partly a function of the number and rate of stimulus presentations (Neimark & Saltzman, 1953; Saltzman, 1954), of the association value of the stimulus (Postman, Adams & Phillips, 1955), of the orienting task (Saltzman, 1953), and of the subjects discriminative and verbal habits (Plenderleith & Postman, 1956).

The aim of this study was to obtain more information on selectivity in incidental learning, and to relate this to the character of the organism's attentional contact with its environment. The design of previous studies has generally precluded this type of investigation. One exception, however, is a study by Stevenson (1954), who set out to examine latent learning in children 3 to 6 years old. Subjects were required to recover a reward from a locked box, the key to which was located among other irrelevant objects in two goal boxes which were positioned at terminal points of a Y maze. One of the irrelevant objects in each of the goal boxes was designated the test object. The key was hidden under, in, or on the test object. Other objects were mere fillers to prevent the test object from becoming unduly conspicuous. Latent learning was considered to be demonstrated if the subject, on recognition of the object, could go to the correct arm of the maze and select the appropriate test object. The results indicated that the ability to locate the object increased significantly with chronological age, and that the frequency of learning was significantly greater when the key was hidden in the test object.

Stevenson's paper raises a general question which is the starting point for the present research. What else, beyond the location of the test object, might the subjects have learned about their environment? Does incidental learning go beyond the test object to the irrelevant features of the environment?

<sup>1</sup>After the initial draft of this paper was accepted for publication, Dr. T. H. Scott was killed in a climbing accident in New Zealand. Therefore, the senior author is alone responsible for final revisions.

In this study, it is proposed: (1) to investigate the limits of incidental learning in adult human subjects in terms of degree of relevance to a task; (2) to test the efficiency of recall and recognition measures in incidental learning; (3) to investigate the effect of heightened motivational level on incidental learning.

#### METHOD

For these experiments a latent learning design of the type used by Spence, Bergmann, and Lippitt (1950) was adopted. In this, S is required to perform a task and it is then established what he has learned about aspects of the environment encountered in doing so.

#### EXPERIMENT I

The task was to find a Yale key concealed under an object in a 6 ft.  $\times$  6 ft. 6 in. cubicle. On entering the cubicle, S found himself beside a shelf; having explored the objects on it his progress was blocked by a screen. Turning left, S faced one end of a closed cupboard running the length of the wall opposite to the shelf; a search among the objects placed along the top of this cupboard carried him past the screen. Turning right again from the cupboard he found objects near him on a table and on top of a large cardboard-box lid lying on the table. Under this lid, when he lifted it aside, were further objects, one of which concealed the key.

#### Subjects

The subjects in this experiment were 57 male undergraduates.

#### Apparatus

The search objects were 20 everyday articles, 10 of which (L) were too small (or wrongly shaped) to conceal the key, while the other 10 (R) could have been key concealers. The key was hidden under a twenty-first article. Four additional objects or features were present which were designated irrelevant objects (I) as they were less likely to conceal the key than L objects. These were a chair against the wall between cupboard and table, a green top to the box lid on the table, a large piece of yellow paper pinned to the table under the box lid, a framed picture on the wall above the table. The search objects were arranged so that they were explored in this order (non-concealers alternating approximately with potential concealers): pen with nib (L), small box lid (R), pencil (L), matchbox (R), (all on the shelf); roll of adding-machine paper (R), small pair of scissors (L), large ink bottle (R), hot-water bottle top (L), piece of candle (L), oil can (R) (all on the cupboard); book (R), small cooking essence bottle (L), cigarette tin (R), narrow chisel (L) (on the table); small spoon (L), fountain pen box (R), small jar (L) (on the large cardboard-box lid); inverted cup (R), box of drawing pins (R), piece of chalk (L), piece of carbon paper concealing the key (under the cardboard-box lid).

#### Procedure

Subjects, tested individually, were instructed as follows: "I want you to find a key just like this (E displays duplicate) which is hidden under an object in that cubicle. When you have found it bring it out to me." On emerging S was asked to recall as many objects as he could remember seeing in the cubicle, and to describe their location. He then picked out from a recognition set (duplicates of the search objects

plus five catch objects) those he recognized and described their location. A verbal recognition test for the additional objects (chair, picture etc.) followed. Finally S was asked questions about his experience and expectations before and during the search. Without his knowledge S's search time was recorded and his progress observed; results from Ss who did not explore the objects in the standard order, or who failed to find the key at a first attempt and went back over the objects, were discarded.

### EXPERIMENT II

#### *Subjects*

Subjects were 20 male undergraduates.

#### *Procedure*

Experiment II was concerned with investigating the effects of increased level of motivation. The procedure was the same as in Experiment I except for the instructions to S which were changed as follows: "I want you to find a key just like this. . . . I want you to find it as quickly as you possibly can, and I am going to time you (*E* indicates stopwatch). When I say "Go," I want you to enter the cubicle and find the key as quickly as you possibly can. go."

### EXPERIMENT III

#### *Subjects*

Subjects were 20 male undergraduates.

#### *Procedure*

Experiment III was concerned with investigating incidental learning under conditions where no set was established by the instructions. In this experiment, therefore, object relevance was not involved. The arrangement of the cubicle was as for Experiment I, except that the objects were made visible without exploration: the cardboard-box lid with its objects on top was placed on the chair and the screen was removed; the book was replaced by a piece of plywood of the same size to prevent S from reading and so being completely distracted from the other objects. On arrival S was asked to wait in the cubicle, being left there for the average search time of the Ss Experiment I (1½ min.); he was then tested and questioned as for Experiment I.

### RESULTS

For purposes of analysis the objects are designated as follows:

Highly relevant (H)—The object which concealed the key

Relevant (R)—The ten search objects which could have concealed the key

Less relevant (L)—The ten search objects too small or wrongly shaped

Irrelevant (I)—Additional objects and features

Together these constitute a tentative relevance gradient from highly relevant to irrelevant. In analysing the results, most attention will be paid to R and L, on which the most complete and best information was obtained. For ease of presentation and comparison, results are shown as percentage frequencies, that is, obtained responses as a proportion of the possible.

TABLE I  
EXPERIMENT I ( $N = 57$ ): PERCENTAGE FREQUENCIES OF  
RESPONSES, ALL CLASSES OF OBJECTS, ON THREE MEASURES  
OF INCIDENTAL LEARNING

Measure	Relevance			
	H	R	L	I
Total recognition	93.0	76.8	36.8	
Correct recognition	93.0	56.1	24.6	15.8
Correct recall	87.7	39.0	14.9	

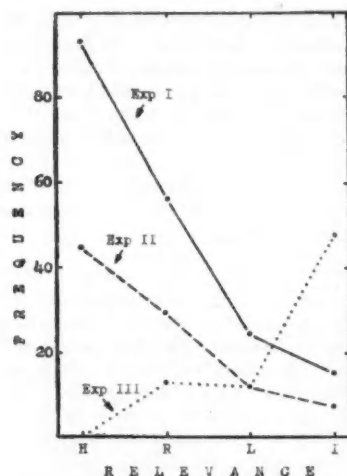


FIGURE 1. Experiments I-III: Percentage frequency of correct recognitions for all classes of objects.

TABLE II  
EXPERIMENT I: ANALYSIS OF VARIANCE OF INCIDENTAL  
LEARNING BY 57 SUBJECTS AT TWO DEGREES OF  
RELEVANCE ON THREE MEASURES

Source	df	MS	F
Relevance	1	437.44	251.40*
Measures	2	127.18	310.19*
Subjects	56	3.87	
R $\times$ S	56	1.74	4.64*
R $\times$ M	2	8.89	34.85*

\* $p = 0.001$ .



*Experiment I*

Percentage frequencies of responses for the four classes of objects on three measures of incidental learning are shown in Table I, and graphically in Figure 1. An analysis of variance on the R and L data (Table II) shows that the effects of the relevance and measures treatments are significant, as are the relevance  $\times$  measures and the relevance  $\times$  subjects interactions. Further analysis showed that significantly more R than L items were recalled ( $t = 5.91$ ,  $p = < 0.001$ ) and recognized ( $t = 6.54$ ,  $p = < 0.001$ ).

*Experiment II*

Percentage frequencies of responses for the four classes of objects on three measures of incidental learning are shown in Table III. An analysis of variance of this data (Table IV) shows that the effects of the relevance and measures treatments are again significant, as is the relevance  $\times$  measures interaction. Further analysis shows that significantly more R than L items are again recalled ( $t = 2.69$ ,  $p = < 0.01$ ) and recognized ( $t = 5.41$ ,  $p = < 0.001$ ). None of the differences between H and R is significant, nor is that between L and I.

TABLE III

EXPERIMENT II ( $N = 20$ ): PERCENTAGE FREQUENCIES OF RESPONSES, ALL CLASSES OF OBJECTS, ON THREE MEASURES OF INCIDENTAL LEARNING (INTENSIFIED MOTIVATION)

Measure	Relevance			
	H	R	L	I
Total recognition	45	50	25.5	
Correct recognition	45	29.5	12.5	7.5
Correct recall	35	20	11.5	

TABLE IV

EXPERIMENT II: ANALYSIS OF VARIANCE OF INCIDENTAL LEARNING BY 20 SUBJECTS AT TWO DEGREES OF RELEVANCE, ON THREE MEASURES

Source	df	MS	F
Relevance	1	41.67	19.47**
Measures	2	26.41	43.30**
Subjects	19	2.03	
R $\times$ S	19	2.14	
R $\times$ M	2	3.21	5.35*

\* $p = 0.01$ .

\*\* $p = 0.001$ .

*Experiment III*

Percentage frequencies of the four classes of objects on two measures of incidental learning are shown in Table V. There was no learning of H, and little difference between R and L objects. I items, however, scored significantly higher (rather than lower) than R ( $t = 7.32, p = < 0.001$ ) or L ( $t = 7.66, p = < 0.001$ ).

TABLE V  
EXPERIMENT III ( $N = 20$ ): PERCENTAGE FREQUENCIES OF  
RESPONSES, ALL CLASSES OF OBJECTS, ON TWO MEASURES  
("CASUAL" MOTIVATION)

Measure	Relevance			
	H	R	L	I
Correct recognition	0	18	17	47.5
Correct recall	0	13.5	12.5	

## DISCUSSION AND CONCLUSIONS

*Incidental Learning and the Relevance Gradient*

The results of this experiment clearly show that an important determinant of what is incidentally learned in a behaviour sequence is the relevance of environmental features to the goal which is directing behaviour. Incidental learning is a function of a "relevance gradient". On the part of the gradient most closely studied (relevant and less relevant objects) the relevant are learned significantly better by every measure in both Experiments I and II. Taken object by object, however, agreement between degree of relevance and amount of learning is not perfect: additional factors probably operate—perhaps primacy and recency effects, and subjective relevance in disagreement with experimentally designated relevance.

Incidental learning, then, is shown to be selective. This selectivity is determined by the set established by a simple goal demand. Compare the results of Experiments I and II with those of Experiment III. In the former, the subject "penetrated" his environment systematically under an immediate goal demand, and subsequently recalled better those objects relevant to this demand. The chances that he will retain incidental information about features of the environment high on the relevance gradient are enhanced, whereas for features low on the gradient the chances are depressed. In Experiment III, however, he entered his environment only "casually." Here H scored lowest (0), and I scored

highest—significantly higher than L or R, which scored about the same. In such casual contact the amount of learning of environment features presumably reflects their intrinsic interest and attention-getting value.

### *Incidental Learning and the Type of Measure*

The results of this study indicate that the amount of incidental learning obtained is partly a function of the measure of learning adopted. A recall test, which is the most selective test of learning, taps mainly relevant objects. A recognition test, a more sensitive test of learning, allows the lower associations of the less relevant objects to appear. These results demonstrate a gradient in the amount incidentally learned from correct recall through correct recognition to total recognition.

### *Incidental Learning under Conditions of Increased Motivation*

One purpose of the present series of experiments was to test the hypothesis that the amount of incidental learning is, within the limits tested, inversely related to the degree of motivation. As can be seen from a comparison of Tables I and III, the increase in motivation which was presumably produced by the instructions given in Experiment II brought about a substantial reduction in retention scores. For example, the amount of learning as revealed by correct recognition falls by about 50 per cent from Experiment I to Experiment II over the whole relevance gradient from H to I. This result is consistent with Bahricks (1954) finding that while intentional learning was faster when the incentive given his subjects was increased, the amount of incidental learning decreased. Similarly Bahricks, Fitts, and Rankin (1952) have shown that increase in incentive leads to a higher degree of selective attention for those parts of a task which the subjects interpret as more important with a resulting tendency to pay less attention to other features of the situation.

In general, the results throw light on the selective agents affecting what and how much is incidentally learned, in a simple, non-repetitive task. The variations in the amount and quality of incidental learning obtained suggest subtle variations in the quality of attention for given features of the organism's environment. The technique employed, and the dependent variable measured, appears to be a flexible instrument which would lend itself to a study of the operation of other variables, for example, age, intelligence, personality variables and clinical conditions, which presumably govern the intricacies of the organism's penetration of his environment. Such study may even contribute to our understanding of the nature and mode of action of such variables in their observable effects on behaviour.

## SUMMARY

Three experiments were carried out to investigate incidental learning in a simple task situation. Ss were required to find a Yale key which was hidden under an object in a cubicle. Half the objects were potential key-concealers and were designated relevant objects (R), the other half being too small or wrongly shaped to conceal the key were designated less relevant objects (L). Additional features in the cubicle were designated irrelevant objects (I).

In Experiment I, 57 Ss carried out a simple task and were subsequently tested for incidental learning on three measures: "correct recall," "correct recognition," and "total recognition." On the three measures, R objects were learned significantly better than L objects.

In Experiment II, 20 Ss performed the same task under intensified motivation. This produced a pattern of incidental learning similar to that obtained under ordinary motivation, but at a much lower over-all level.

In Experiment III, 20 Ss were exposed "casually" to the same situation, being required only to wait there. Their incidental learning showed none of the patterns obtained in the earlier experiments; a different pattern based presumably on the intrinsic interest of the environmental features was obtained.

The results are explained in terms of a grading of attention by central processes, affecting the character of the organism's contact with specific aspects of the material environment.

## REFERENCES

- BAHRICK, H. P. Incidental learning under two incentive conditions. *J. exp. Psychol.*, 1954, 47, 170-172.
- BAHRICK, H. P., FITTS, P. M., AND RANKIN, R. R. Effect of incentives upon reactions to peripheral stimuli. *J. exp. Psychol.*, 1952, 44, 400-406.
- JENKINS, J. G. Instruction as a factor in incidental learning. *Amer. J. Psychol.*, 1933, 45, 471-477.
- LEPLEY, W. M. A gradient in incidental learning. *J. exp. Psychol.*, 1935, 18, 195-201.
- NEDMARK, E., & SALTZMAN, I. J. Intentional and incidental learning with different rates of stimulus presentation. *Amer. J. Psychol.*, 1953, 66, 618-621.
- PLENDERLEITH, M., & POSTMAN, L. Discriminative and verbal habits in incidental learning. *Amer. J. Psychol.*, 1956, 69, 236-243.
- POSTMAN, L., & SENDERS, V. L. Incidental learning and the generality of set. *J. exp. Psychol.*, 1946, 36, 153-165.
- POSTMAN, L., ADAMS, P. A., & PHILLIPS, L. W. Studies in incidental learning: II. The effects of association value and of the method of testing. *J. exp. Psychol.*, 1955, 49, 1-10.
- SALTZMAN, I. J. The orienting task in incidental and intentional learning. *Amer. J. Psychol.*, 1953, 66, 593-597.
- SALTZMAN, I. J. Comparisons of incidental and intentional learning after different numbers of stimulus presentations. *Amer. J. Psychol.*, 1954, 67, 521-524.
- SPENCE, K. W., BERGMANN, C. AND LIPPITT, R. A. A study of simple learning under irrelevant motivational-reward conditions. *J. exp. Psychol.*, 1950, 40, 539-551.
- STEVENSON, H. W. Latent learning in children. *J. exp. Psychol.*, 1954, 47, 17-21.

## TACTUAL AND VISUAL INTERPOLATION: A CROSS-MODAL COMPARISON<sup>1</sup>

A. V. CHURCHILL

*Defence Research Medical Laboratories, Toronto*

IN A STUDY of tactual<sup>2</sup> and visual interpolation (Churchill, 1959), subjects were required to interpolate to tenths of the range between two brass reference rods,  $\frac{1}{2}$  and  $1\frac{1}{2}$  in. in diameter, designated as "zero" and "ten" respectively. The nine interpolated positions were represented by a series of intermediate rods of regularly increasing diameters. It was demonstrated that subjects underestimated the diameter of the intermediate rods tactually, that is, "felt" the rods to be smaller in diameter than they actually were, and overestimated the diameter of the intermediate rods visually, that is, "saw" the rods to be larger in diameter than they actually were.<sup>3</sup>

If rod diameter, *per se*, were underestimated tactually, then the diameter of the reference rods, "zero" and "ten," would be underestimated also, with a resultant absence of directional bias, that is, the direction of error would be constant for all rods, both intermediate and reference. If rod diameter, *per se*, were overestimated visually, then the reference rods also would be overestimated, again with a resultant absence of directional bias. Since opposed directional biases of the errors were evident in the experiment reported earlier, it appears that under both sensory conditions the intermediate and reference rods were perceived differently, that is, under the condition of tactual stimulation the intermediate rods were underestimated and/or the reference rods were overestimated; the reverse being true under the condition of visual stimulation.

The present paper reports a study of these directional biases, in which the intermediate rods were presented to one sensory modality and the reference rods to the other. In addition, the tactual and visual conditions of the earlier study were repeated. The four stimulus conditions were designated as: T, tactual presentation of both intermediate and reference rods; Ti, tactual presentation of the intermediate rods and visual presentation of the reference rods; V, visual presentation of both intermediate and reference rods; Vi, visual presentation of the intermediate

<sup>1</sup>DRML Report no. 164-11, PCC no. D77-94-20-27, H.R. no. 190.

<sup>2</sup>The term "tactual" is used here to simplify presentation, and is defined by the task.

<sup>3</sup>An error of omission in the earlier report (Churchill, 1959) can be corrected here. The Ss were members of the laboratory workshops staff, and all were right-handed.

rods and tactual presentation of the reference rods.  $T_i$  and  $V_i$  are referred to hereafter as cross-modal comparisons.

If the bias were associated with the intermediate rods, then, under cross-modal stimulation, the direction of the bias would be dependent upon the modality stimulated by the intermediate rods, that is, the bias would be in the same direction as when both the intermediate and the reference rods were presented to that modality. If, on the other hand, the bias were associated with the reference rods, then the direction of the bias would be dependent upon the modality stimulated by the reference rods. In the event that the intermediate and reference rods contributed equally to the directional bias, but in opposite directions, then the bias would tend to disappear.

### METHOD

#### *Apparatus*

The apparatus has been described (Churchill, 1959). The reference rods were  $\frac{1}{8}$  and  $1\frac{1}{8}$  in. in diameter and were designated as "zero" and "ten" respectively, with the "zero" reference rod placed 12 in. to the left of the "ten" reference rod. The nine intermediate rods, representing the interpolated positions "one" through "nine," increased in diameter by  $\frac{1}{16}$  in. increments, that is, the rod representing the interpolated position "one" was  $\frac{9}{16}$  in. in diameter, that representing "two" was  $\frac{10}{16}$  in. in diameter, etc. Interchangeable panels, mounted in front of the apparatus, permitted the presentation of each of the four stimulus conditions.

#### *Procedure*

Seven of the laboratory workshops personnel used in the previous study served as  $S$ s for the present experiment. Subject was seated in front of the apparatus with his hands and the tactually sensed rods, reference and/or intermediate, screened from his view by a drop-cloth. In this position the visually sensed rods were presented at a viewing distance of approximately 28 in. Subject was permitted continuous "sensing" of the reference rods under all experimental conditions and was not restricted as to time or number of contacts with the intermediate rods. Each of the nine intermediate rods was presented twice in a random series of 18 presentations; three such series were presented to  $S$  before changing conditions. The four experimental conditions were presented in random order.

Subjects' task, under all stimulus conditions, was to report the fraction of the interval between the reference rods represented by the intermediate rod on each presentation; for example, for the intermediate rod  $\frac{1}{8}$  in. in diameter the correct response would be "four." Subjects were restricted to the nine responses "one" through "nine."

### RESULTS

Errors of interpolation were classified as overestimations or underestimations, that is, positive errors or negative errors. Since error magnitude does not alter the main results, only error frequency data are presented here (the exceptions are data on constant errors where error magnitude is included).

Table I shows the percentages of interpolations overestimated and underestimated for each of the four experimental conditions. Comparative percentages for conditions T and V are shown, in parentheses, for the same subjects in the earlier study.

TABLE I  
PERCENTAGES OF INTERPOLATIONS OVERESTIMATED AND  
UNDERESTIMATED UNDER FOUR EXPERIMENTAL CONDITIONS  
( $N = 7$ )

Condition	Over- estimation	Under- estimation	Total error
T	17.5 (14.3)*	49.7 (42.3)*	67.2 (56.6)*
Ti	21.4	46.8	68.2
V	48.2 (51.9)*	18.5 (17.7)*	66.7 (69.6)*
Vi	54.2	14.8	69.0

\*Results obtained by same Ss, 10 months earlier (see Churchill, 1959).

From Table I it is apparent that the total percentage of interpolations in error was essentially the same for the four experimental conditions. As in the previous study, subjects underestimated the intermediate rods tactually and overestimated them visually. For the cross-modal condition Ti, it is seen from Table I that the tactual presentation of the intermediate rods produced an underestimation bias as did condition T, the tactual presentation of both the intermediate and the reference rods. For the cross-modal condition Vi, the visual presentation of the intermediate rods produced an overestimation bias as did condition V, the visual presentation of both intermediate and reference rods. Apparently it makes little difference to which sensory modality the reference rods are presented; the modality to which the intermediate rods are presented is primarily responsible for the direction of the error bias.

TABLE II  
ANALYSIS OF VARIANCE OF FOUR  
EXPERIMENTAL CONDITIONS ( $N = 7$ )

Source	df	MS
S (Subjects)	6	267.67*
C (Conditions)	3	7.33
T (Trials)	2	169.03
S $\times$ C	18	168.62*
S $\times$ T	12	53.16
C $\times$ T	6	57.95
S $\times$ C $\times$ T	36	38.84
TOTAL	83	

\* $p < 0.01$ . Theoretical residual 45.6.

The data were transformed to degrees ( $\theta = \sin^{-1}/p$ ) to satisfy the assumptions of analysis of variance (Quenouille, 1950). Results of the analysis of the four stimulus conditions are presented in Table II.

Table II shows an error term of the same order of magnitude as the theoretical residual (Walker & Lev, 1953, p. 423) indicating that the performance of subjects was consistent from trial to trial. The variability indicated by the inflated  $S \times C$  interaction was found, from a separate analysis, to be largely contributed by one person. Exclusion of this subject resulted in an  $S \times C$  interaction which was barely significant at the .05 level of confidence.

The constant error was calculated for each of the nine interpolated positions, and the four stimulus conditions. These data are shown in Figure 1.

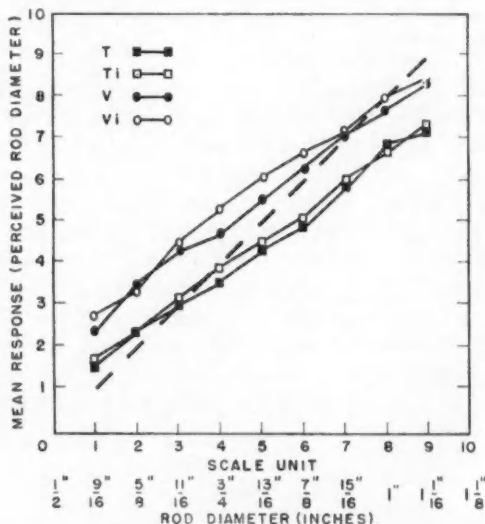


FIGURE 1. Mean constant error for each interpolation position under four stimulus conditions ( $N = 7$ ). (The broken diagonal line represents "zero" constant error.)

From Figure 1, it is seen that the differences within the two related pairs of curves,  $T$  versus  $T_i$ , and  $V$  versus  $V_i$ , are small compared to the over-all difference between the pairs of curves,  $T$  and  $T_i$  versus  $V$  and  $V_i$ . These data lend additional support to the conclusion, drawn above, that



the intermediate rods are primarily responsible for the directional biases of the errors.

The percentages of interpolations overestimated and underestimated were calculated for each of the nine intermediate rods and the four stimulus conditions. These data have been plotted in Figure 2 (A and B).

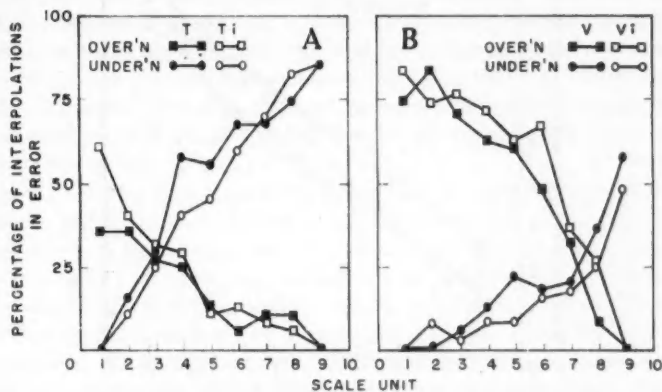


FIGURE 2. Percentages of overestimations and underestimations for each interpolation position, under four stimulus conditions ( $N = 7$ ).

The preponderance of underestimations under conditions T and Ti is apparent from Figure 2A, as is the preponderance of overestimations under conditions V and Vi from Figure 2B. Of greater interest, however, is the similarity of the error patterns, in terms of both percentage and direction, across the nine interpolated positions for both T and Ti, and for V and Vi. Note that the pattern for T and Ti is almost a mirror-image of that for V and Vi. It should be noted also that a central tendency, or bias of error towards the midpoint of the scale, is demonstrated by the decrease in overestimations and the increase in underestimations as the rod diameter increases, under all four conditions (cf. Hollingworth, 1910).

Figure 3 shows the constant errors obtained under conditions T and V in the present experiment, and those obtained from the same subjects in the experiment reported earlier.

It is apparent from Figure 3 that there was little difference, under conditions T and V, between the results of the two studies conducted ten months apart. The consistency of the results obtained from the two studies was tested by analysis of variance. The analysis showed no differences in the main effects or in the interactions, indicating that subjects

performed this particular interpolation task with equal accuracy under tactual or visual conditions. The analysis also showed a high level of reproducibility of the results obtained under these stimulus conditions.

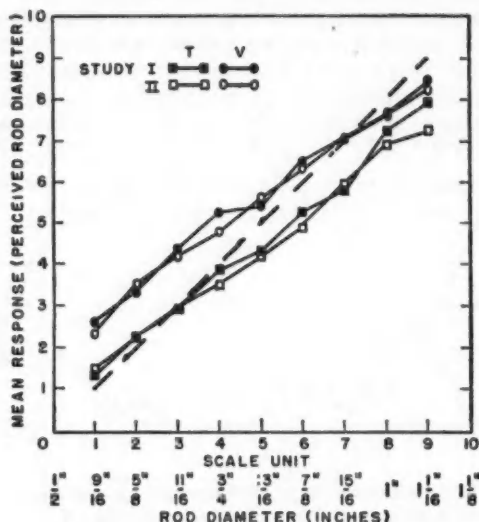


FIGURE 3. Mean constant error for each interpolation position under two stimulus conditions for two studies ( $N = 7$ ). (The broken diagonal line represents "zero" constant error.

### DISCUSSION

The procedure employed in the present study required subjects to sense the smaller, or "zero," reference rod with the left hand and the larger, or "ten," reference rod with the right hand. When tactually judging the intermediate rods, subjects were permitted to use either hand, or each in turn. In general, subjects tended to compare the rod to be judged with each reference rod in turn, that is, grasping it with the right hand to compare it with "zero," and then with the left hand to compare it with "ten." The absence of major differences in the results obtained under conditions T and Ti, that is, tactual judgment of the intermediate rods and tactual (condition T) or visual (condition Ti) sensing of the reference rods, indicates that the procedure did not seriously bias the data. Further justification for this procedure is given by Kelvin (1954) who reported an "absence of lateral dominance" between the hands (p. 31).

As shown in Table I, and verified by analysis of variance, there were no over-all differences in the accuracy with which size judgments were made under the four experimental conditions. These results are in agreement with those reported by Kelvin (1954). His subjects were presented with metal disks, varying in diameter from 57.5 to 62.5 mm., which were judged either larger or smaller than a standard disk 60 mm. in diameter. Kelvin explained these results in terms of a modified concept of Head's (1920) "Schema." Results of a more recent study, Kelvin and Mulik (1958), in which the standard disk was not at the midpoint of the series, showed cross-modal discrimination to be less accurate than discrimination within one sense ( $p < .05$ ). From these results the authors conclude that "co-ordination of Sight and Touch cannot be explained in terms of the schema postulated earlier" (p. 191).

Jastrow (1886) investigated the perception of extension in conjunction with the use of the eye, the hand, and the two senses. A line of a given length, from 5 to 120 mm., was fixated by the eye and the subject was required to (a) select one of a visually presented series of lines which he judged to be equal to the standard length, or (b) set a distance between the thumb and forefinger, on a movable carriage, which he judged to be equal to the standard length. This procedure was followed for the other two conditions, that is, the presentation of the standard length to the hand, and the judgment made by hand or eye. Contrary to the findings reported in the present study, Jastrow found that subjects judged extension more accurately by eye than by hand, and more accurately by one sense than across two senses. He concluded that judgments within one sense involved one brain centre, making the operation easy and the error small; whereas, judgments across two senses involved two brain centres, making the operation difficult and the error large.

The "central tendency" of judgment, demonstrated by the decrease in overestimations and the increase in underestimations which accompanied the increase in rod diameter, under all four experimental conditions (Figure 2), is in apparent disagreement with the results reported by Jastrow. When his subjects fixated the standard length with the eye, they underestimated short lengths and overestimated long lengths when judgments were made by eye, and overestimated all lengths when judgments were made by hand. When the standard length was sensed with the hand, subjects overestimated short lengths and underestimated long lengths (central tendency) when judgments were made by hand, and underestimated all lengths when judgments were made by eye. The absence of a "central tendency" effect under three of these of four conditions is probably due, in part, to the method of presenting the results. Hollingworth (1910), and Watson (1957), have shown the "central tendency" effect to be operative underneath strong constant error biases.

Analysis of the data also revealed a high level of reproducibility of the results obtained under the tactual (T) and visual (V) stimulus conditions (since cross-modal conditions were not included in the earlier study, a measure of the reproducibility of the cross-modal results is not available).

Generalization of the findings reported here is limited by the subject population, that is, laboratory workshops personnel.

### SUMMARY

The present experiment had a twofold purpose, (a) to study the directional bias of errors when making cross-modal judgments of size, and (b) to establish the consistency of results obtained for both tactual and visual judgments in an earlier experiment. Ss were required to judge the position of each of nine rods, varying in diameter by a constant increment, between two reference rods which were designated as "zero" and "ten." The intermediate rods were judged tactually or visually while S sensed the reference rods tactually or visually. Thus, judgments were made under four different conditions.

The results indicate that this task is performed with equal accuracy under all four experimental conditions. The data also reveal that the directional bias of the errors is dependent on the sense stimulated by the intermediate rods, that is, an underestimation bias when the intermediate rods are judged tactually whether the reference rods are sensed tactually or visually, and an overestimation bias when the intermediate rods are judged visually whether the reference rods are sensed visually or tactually. The "central tendency" effect obtains under all four experimental conditions. Results are consistent with those reported in an earlier study.

### REFERENCES

- CHURCHILL, A. V. A comparison of tactual and visual interpolation. *Canad. J. Psychol.*, 1959, 13, 23-27.
- HEAD, H. *Studies in Neurology*. Oxford Univer. Press, 1920.
- HOLLINGWORTH, H. L. The central tendency of judgment. *J. Phil., Psychol., & Sci. Methods*, 1910, 7, 461-469.
- JASTROW, J. Perception of space by disparate senses. *Mind*, 1886, 11, 539-554.
- KELVIN, R. P. Discrimination of size by sight and touch. *Quart. J. exp. Psychol.*, 1954, 6, 23-34.
- KELVIN, R. P., & MULIK, A. Discrimination of length by sight and touch. *Quart. J. exp. Psychol.*, 1958, 10, 187-192.
- QUENOUILLE, M. H. *Introductory Statistics*. London: Butterworth-Springer, 1950.
- WALKER, HELEN M., & LEV, J. *Statistical Inference*. New York: Holt, 1953.
- WATSON, W. A. Contrast, assimilation, and the effect of central tendency. *Amer. J. Psychol.*, 1957, 70, 560-568.

## CHILDREN'S UNDERSTANDING OF NUMBER AND RELATED CONCEPTS<sup>1</sup>

P. C. DODWELL  
*Queen's University*

SOME YEARS AGO an English translation of Piaget's "*La Genèse du nombre chez l'enfant*" was published (Piaget, 1952a). Unlike some of his earlier work on children's language and thinking (e.g., Piaget, 1926; 1929), this study of number concepts and children's ability to use numerical operations has not produced very great interest, acclaim, or criticism; nor, to judge by published reports, has it stimulated much research by independent investigators. This is surprising, since the work is an improvement on Piaget's earlier studies in at least two respects. First, the theoretical background is much more precise, but at the same time more elaborate than his earlier theories of cognitive development, and secondly the empirical investigations are more objective, described in sufficient detail to be essentially repeatable, and not open to the earlier criticisms of too heavy a reliance on interpretation of verbal statements and the possibility of "projecting" the experimenter's ideas into these interpretations (Hazlitt, 1930; McCarthy, 1930). This is not to say that the later experiments are above criticism; Piaget usually fails to specify the number of subjects used in any one investigation, frequently bases a generalization (apparently) on the behaviour of but one or two children, and does not say whether a particular type of observed behaviour is universal, typical, or merely found occasionally, at any particular age or stage of development.

One study (Estes, 1956) reports findings which claim to refute completely Piaget's statements about young children's responses to problems involving numerical concepts. It would appear that this refutation should be taken seriously since the test situations used were, ostensibly, the very situations which Piaget himself describes. Even though the number of subjects used was not too small (52), the generally inimical tone of the paper tends to raise a doubt in the reader's mind, especially since it appears that Estes was not familiar with the main body of Piaget's work on number concepts. Having supervised several small-scale projects (unpublished) which on the whole tended to support Piaget's state-

<sup>1</sup>This investigation was supported by a grant from the Arts Research Committee of Queen's University, whose assistance is gratefully acknowledged. Thanks are due also to students in the Department of Psychology who took part in the investigation, and to the Inspector of Public Schools in Kingston and the teachers who made it possible.

ments, I felt that a more thorough investigation was called for. Such an investigation, based on study of some 250 children, is reported below.

### PIAGET'S THEORY

Piaget's theory of the development of number concepts is a particular application of his general theory of intelligence (Piaget, 1950). The central idea is that rational behaviour, and in particular the production of rational (operational) solutions to problems involving number, develop from a more primitive form of thinking which is syncretistic and egocentric; that is, which does not operate with categories and relations which are well defined, articulated, and self-consistent, and also does not apply rules which are independent of the "viewpoint" of the operator. Piaget considers this type of thought as being not merely a poor attempt at the thinking which is characteristic of the rational adult, but rather as having positive properties of its own, which are both characteristic for a particular stage of cognitive development, and limit the type of understanding which is possible at that stage. Piaget describes a number of stages of cognitive development; the present study is concerned with but three of them. At about the sixth year, when most children display a spontaneous interest in numbers, and have already learned to count, Piaget claims that they have only a vague notion of what the concept of "number" is. This can be demonstrated in a variety of ways, some of which will be described below. At this stage judgments about problems involving numbers are to a large extent determined by what the child *perceives*, so that if a perceived configuration is changed, the numerical judgments made about the situation will be likely to change. The same considerations apply to judgments about quantities; quantity for a young child turns out to be determined by perceived characteristics, rather than a logical notion which obeys laws of conservation, etc. This stage, which is clearly egocentric (bound to a particular perceptual point of view), is called by Piaget the stage of "global comparisons."<sup>2</sup> It is followed, according to Piaget, by an "intuitive" stage in which the child starts to realize that judgments of quantity and number cannot be made simply in terms of perceived attributes; it starts to grasp, fleetingly and unclearly, that quantity and number are attributes of objects, or sets of objects, which remain invariant under perceptual transformations. One might say that the child starts to emancipate itself from the purely perceptual field, although its cognitive activity is still bound to judgments about objects in the perceptual field. The third stage,

<sup>2</sup>The labelling of stages here used is the standard nomenclature of the English translation of *The Child's Conception of Number* (Piaget, 1952).

in which judgment becomes completely "operational"—no longer bound to perceived patterns, and not egocentric—is called by Piaget the stage of "concrete operations." Operations and judgments start to manifest stability, self-consistency, and "reversibility,"<sup>3</sup> but can still only be performed on perceived objects, not in the abstract.

Specifically, operations which are necessary conditions of an understanding of numbers are, according to Piaget, the ability to deal with the equivalence of cardinal classes in terms of one-to-one correspondence, and the ability to deal with transitive relations such as "greater than" and "less than." This suggests that in order to understand what a number is, a child must be able to manipulate and make judgments about perceived objects in such a way that (a) the order, or perceived pattern, of elements in a group of objects does not influence judgments about the number of objects present, and (b) the child should be able to arrange objects in series according to some obvious criterion such as size, and should be able to deal with ordinal correspondence between different series (i.e., judgments involving relative position in the series). This is, according to Piaget, the psychological parallel of the fact that number, as a concept, can be reduced to the logically more primitive notions of order, one-to-one correspondence, etc., as Frege, Russell and others have shown (Russell, 1919).

The second (intuitive) and third (operational) stages are held to occur for most children in about the seventh and eighth years, although individual variations are considerable. Piaget has a good deal to say about the transition from one stage to another, but the details of the transitions need not be elaborated here: more complete descriptions of Piaget's theories are to be found in his own writings, and in an article of the present writer (Piaget, 1950, 1952a; Dodwell, 1957).

#### PIAGET'S EVIDENCE

The evidence Piaget produces consists mainly in demonstrations that young children make inconsistent judgments about numerosity, even when they can count, and that the attainment of consistency follows on the "realization" of the nature of one-to-one correspondence and ordinal equivalence in series.

Weaknesses in Piaget's evidence have been pointed out above; further objections can be raised, on the ground that Piaget does not indicate

<sup>3</sup>An operation is reversible, in Piaget's terminology, if the child understands that it has an inverse operation which cancels the original one; for example, "add two" has the inverse "subtract two," and if a child understands this, the operation of addition is said to be reversible for the child.



how *consistent* children are in the types of response they make, nor does he put forward evidence which shows that the stages follow each other in the order required by the theory, in all children.

The test situations used by Piaget involve different types of material (beakers and liquid, counters, eggs and eggcups, dolls, sticks, cards of different sizes, etc.) and all involve the subjects in making judgments about quantities or numbers, and usually also involve manipulation of the materials. A catalogue of these situations will not be given here; the test situations and techniques used in the present investigation were taken over with comparatively little modification from Piaget's work, and their description below will serve to illustrate the sorts of evidence on which Piaget's claims are based.

### PROCEDURE

#### *Aim*

The aim was to assess the generality of the types of behaviour described by Piaget for children between the ages of about 5 and 8 years old, to examine age trends, the consistency of behaviour at any particular age, and to assess these factors as evidence for a theory of cognitive development in terms of the three stages described above.

#### *Method and Materials*

Five persons took part in the investigation, four of them as testers, three (including two testers) as scorers. The subjects—250 of them—were all children in Kingston public schools. They were tested individually with the test to be described below. They were in five different schools, were all in Kindergarten, Grade I or II, and their ages ranged from 5;1 to 10;1 years. No I.Q.'s were available for kindergarten children, but ratings of "above average," "average," and "below average" were obtained from the teachers. For Grade I and II pupils, I.Q.'s were available, measured on a group test; an attempt was made to have the I.Q. proportions in each grade reflect population proportions, at least approximately. No very careful fitting was attempted, for two reasons: first, the number of children available for selection was limited, and secondly, the I.Q.'s available were not particularly accurate measures. In the samples the means were somewhat too heavily represented, at the expense of the extremes, as Table I demonstrates. Since the aim of the investigation is to assess the generality of Piaget's findings for "normal" children, this over-representation of the middle categories is hardly a serious defect. The age distributions within each grade were not symmetrical: the distribution for kindergarten was—as one might expect—positively skewed, with a range from 5;2 to 6;8 and mode of 5;5. The Grade I age distribution was also positively skewed, but to a smaller extent, and had a larger range (6;0 to 8;8), with the mode at 6;9. Grade II ages were fairly evenly distributed in the range 7;3 to 8;7, plus a few stragglers at the upper end, with no clear mode. With the numbers of children available it was not possible to match age and I.Q. to obtain symmetrical I.Q. distributions at age, rather than at grade levels. The five schools were in different parts of the city, and all socio-economic levels were represented.

In Piaget's method of investigation the experimenter is to a great extent free to question the child, follow up ambiguous replies and so on, as the situation demands; the argument in favour of this method is that it is the only way of gaining a com-

TABLE I  
DISTRIBUTIONS OF I.Q. IN SAMPLE

	I.Q.				
	-85	86-95	96-105	106-115	116+
Population I.Q. proportions*	0.16	0.21	0.26	0.21	0.16
Sample I.Q. proportions					
(Grade I, $N = 110$ )	0.05	0.26	0.34	0.28	0.08
(Grade II, $N = 55$ )	0.07	0.24	0.33	0.25	0.1
Kindergarten ( $N = 85$ )					
Below average: 0.23					
Average: 0.53					
Above average: 0.24					

\*Population I.Q. proportions are calculated on a normal distribution with mean of 100, standard deviation of 15.

prehensive insight into the mental processes of the child. One can object to it on the grounds that, however careful the experimenter, different sorts of question, and different sequences of questions, may influence the sorts of reply given and the behaviour manifested, and hence raise the variability of behaviour. This is not desirable if one wishes to compare performances of children at different ages, estimate norms, etc. For this reason the form of the present investigation was kept fairly standardized. Five different types of test material were used, four of them very similar to materials used by Piaget (see Table II).

TABLE II  
TEST SUBGROUPS: SITUATIONS AND MATERIALS\*

Subgroup	Situation	Materials
I	Relation of perceived size to number	Beakers and beads
II	Provoked correspondence	Eggs and eggcups
III	Unprovoked correspondence	Red and blue poker chips
IV	Seriation	Dolls and canes of graded size
V	Cardination and ordination	Wooden cubes and doll

\*For description of procedure, see text.

First, the child was asked to count out 12 beads, picking up two at a time (one in each hand) and placing them in two similar glass beakers, six beads in each. It was asked, "Are there the same number of beads in each glass?" and then, "How do you know?" Answers were recorded on a standard test blank. Only three answers are possible to the first question (Yes, No, Don't know), but a number of different answers could be given to the second. The most common answers were: "I counted the beads out," "the beads look the same," "they are the same height in the two beakers," "Don't know." If an answer substantially different from one of these was given, it was written out; otherwise the investigator ticked a box on the test blank for the relevant answer. The beads from one beaker were then poured into a taller, narrower beaker, and the child was asked: "Are there the same number of beads in each glass now?", and if the child answered "No" it was asked: "Which has more?" and "Why?" The most frequent answers to the last question were: "This looks

higher," "This looks more," or "Don't know," and again if a substantially different answer was given, it was written out. The child was then asked to count out eight pairs of beads into dissimilar beakers, and was asked: "Are there the same number in each glass?" Whether the child answered yes or no, it was asked why it had judged the two numbers to be the same or different. Again common categories of answers were listed on the test blank; other answers were written down. The beads from the narrower beaker were then poured into a beaker similar to the first, and the child was asked again to say whether there were the same number in each, and responses were recorded as before.

The second situation involved what Piaget calls "provoked correspondence," in this case correspondence between a set of eggs and eggcups ("provoked" because there is an obvious perceptual—and utilitarian—relation between the egg and the cup). Six eggcups were placed in a row and the child was asked to put one egg in each cup, then asked: "Are there the same number of eggs and cups?" with if necessary, subsidiary questions. The eggs were then laid out in front of their cups, thus maintaining a clear perceptual relation between the two sets. The child was again questioned about the number of eggs and cups. The eggs were then bunched up together, destroying the perceptual correspondence, and the child was questioned. Whether the child said there were or were not the same number of eggs and cups, it was asked how it arrived at its answer. Finally, the child was asked to replace the eggs in the cups, and asked if there were now equal numbers again. Children who had thought the bunched up eggs different in number from those in the cups were asked to reconsider their decision ("So were there as many eggs as cups when the eggs were all bunched up here? . . . Why not?").

The third situation involved "unprovoked correspondence." The experimenter laid out six blue poker chips in a row, gave the child a box of red chips and said: "Can you put out another row like this one?" Nearly all children put out six red chips in a similar row close to the originals. They were then asked to count the numbers in the two rows, and to say whether both rows had the same number of pieces. The blue chips were then spread out, but still in a straight line, and the child was asked: "Which row has more pieces in it now?" (One leading question out of more than 40 was held to be not excessive. It would of course be interesting to find out what effect the form of the question has on the judgments given.) Whatever the answer, the child was asked: "Can you make as many in your row as there are here?" The almost universal response, even among those who thought the rows of unequal length contained equal numbers, was to shift the red counters out, thus reforming the perceptual correspondence between the sets. Some of the younger children added more pieces to their own row, until the two rows were of equal length. The blue pieces were then put back in their original positions, and the questions were repeated: "Are there the same number in each row now?" and "Can you make the rows the same again, with the same number of pieces in each?"

The fourth situation involved judgments about series of objects of different sizes. The investigator placed eight plywood dolls on the table, in a line from smallest to largest, explaining to the child what he was doing. He then produced some canes, also of different sizes, and placed the two smallest in line with the two smallest dolls, but closer to the child, with appropriate verbal explanation: "Each man has a stick to walk with; a small one for the smallest man, a larger one for the next. . . . I want you to put down the stick which belongs to each man in front of him." The number correctly placed by the child was recorded. The row of dolls was spread out, and the child asked which stick belonged to the smallest doll; the second from largest; the fourth from smallest. The dolls were not mentioned, but pointed to, as the

experimenter asked his questions. The child was then asked how he had decided which stick belonged to a particular doll. Answers were recorded in the same way as before.

The fifth and final situation involved building a "staircase" with 2 in. wooden cubes; finding out what the child understands of the process should elucidate its grasp of the relation between ordinal and cardinal numbers. Piaget made use of a similar situation, except that he used rectangular cards of different shapes which could be laid flat on a table to form a "staircase." It was felt that, apart from providing a more realistic situation, the relationship between ordinal and cardinal properties (first—one; second—two, etc.) would be clearer if each class (column of blocks forming a stair) could be seen as composed of individual and separate blocks. The investigator demonstrated by building the first two stairs, and asked the child to build the next one. Then he asked: "Which stair is it? . . . This is the first, this is the second, and this is the . . .?" The child was then asked to build the next two steps, being helped if necessary. It was asked: "What would be the next step? . . . How many blocks would there be in it?" Then, pointing to the third stair: "How many steps will the man have climbed to get here? . . . How many has he still to go?" The third step was removed, and the child was asked how many blocks it had contained, then it was asked to build that stair, but not in its position in the staircase. The last three questions, of a more abstract and hypothetical character were: "If I built 10 steps, how many blocks would there be in the highest step?" "What would the next step be called?" and "How many blocks would there be in it?"

Altogether, 54 questions were listed on the test blanks as "standard questions,"<sup>4</sup> and were always used except in a few cases. For instance, children who obviously knew that the number of eggs and cups did not change when the eggs were bunched up were not all asked the subsidiary questions about how they arrived at their judgments. On the other hand "non-standard" subsidiary questions were occasionally introduced if a child's response had been ambiguous. A small degree of flexibility was thought to be desirable in the administration of the test, despite the arguments mentioned above in favour of standardization.

## RESULTS

It was found that, for all subgroups of the test, some children gave answers which were "non-operational" in Piaget's sense, even though they had some idea of what a number is, would count, and appeared to understand the questions. Thus, many children, after counting equal numbers of beads into dissimilar beakers would judge, on *looking* at the beakers, that the numbers in each were not the same. Similarly, in manipulating the chips, judgments about number were frequently changed when the perceived configuration changed, although such children might realize that counting was an operation relevant to their judgments. A distinction could be drawn between children who only made judgments in terms of what was perceived (Piaget's first stage) and those who, whilst realizing the inadequacy of judgments in this vein, were not able to apply

<sup>4</sup>It was decided not to spell out all 54 questions in this paper; anyone interested in a complete record can obtain one of the test blanks by writing to the author.

the operations of counting, ordering, etc., in a consistent fashion (Piaget's second stage). These children could again be distinguished from those in the third, fully operational stage. The methods used for discriminating the stages more precisely are given below.

The test blanks were scored in two ways. First, a simple point score for the number of questions answered correctly and number of items correctly performed was obtained. This measure, as one might expect, is positively correlated with age; the correlation scatter plot is shown in Figure 1. Table III shows various correlations and partial correlations of this score with other measures.

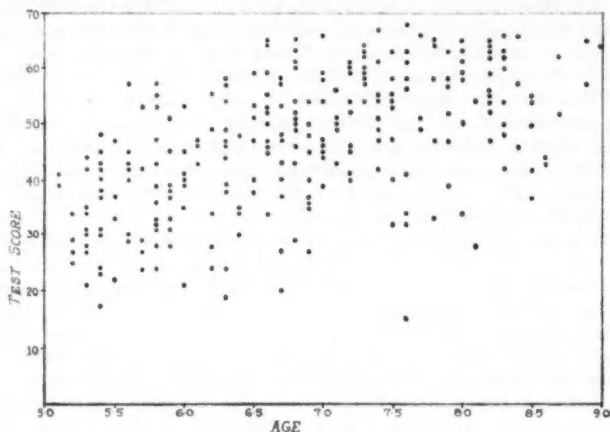


FIGURE 1. Correlation scatter plot of age and point score; Kindergarten, Grades I and II.

TABLE III  
CORRELATIONS OF POINT SCORE AND A SCORE  
WITH MEASURES

	Point score	A score
Age	0.52	-0.56
I.Q.*	0.28	-0.24
A score	-0.71	—
Age* (I.Q. constant)	0.62	-0.63
I.Q.* (Age constant)	0.44	-0.44

\*Grades I and II only,  $N = 165$ .

All correlations significant,  $p < 0.005$  or better.

A more important measure, as far as the present investigation is concerned, is the "A score." For convenience, Piaget's three stages—of global comparisons, intuitive judgments, and concrete operations—were labelled

A, B, and C. Any response made which indicated that a child was in stage A was called an "A response," and the "A score" was simply the number of A responses made. The highest possible number of A responses was ten. In arriving at this number, the following factors were considered: first, only one A response was possible for any one test situation, so that if a child said twice in the same situation that he had judged numbers of beads by their relative heights in the beakers, this was considered one A response. If, however, the child said this in two different situations involving the same material (e.g., when the beads were poured into dissimilar beakers, and when they were counted into dissimilar beakers), it was counted as two A responses. An A response was one which clearly indicated that the child was basing its judgment on perceived characteristics alone, and this criterion was applied without difficulty, with agreement between different scorers. The A score is obviously not independent of the point score, and one expects a negative correlation between them. Table III shows correlations between A score and various other measures,

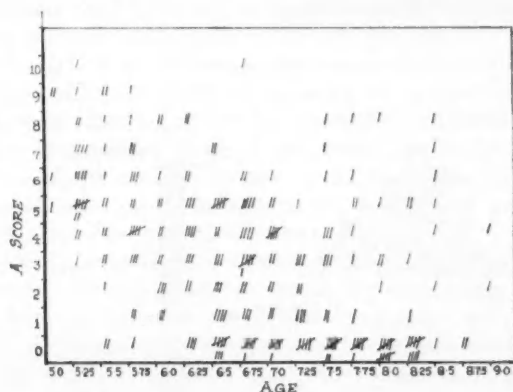
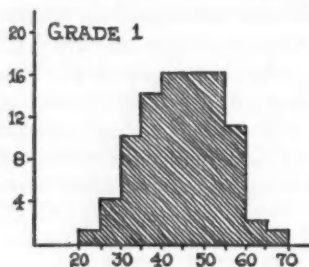
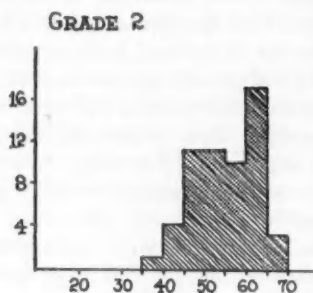
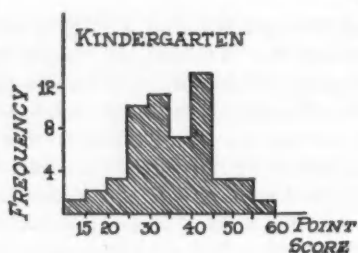


FIGURE 2. Correlation scatter plot of age and A score: Kindergarten, Grades I and II. (Age scale shows lower limit of each interval.)

Figure 2 the correlation scatter plot for A score and age. As with the point score, we find that both chronological age and I.Q. are covariants of the A score. The associations between grade, point score, and A score are illustrated graphically in Figure 3.

Unfortunately, the investigation of the effects of length of schooling on ability to deal with number concepts was beyond the scope of the pre-

A



B

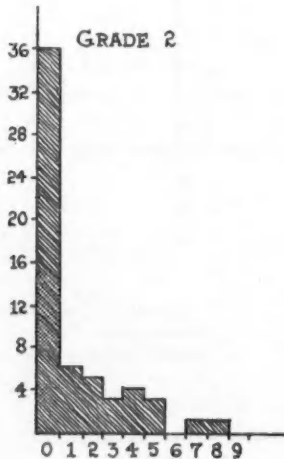
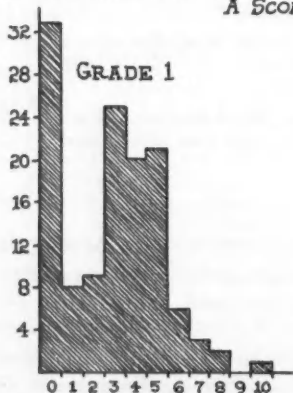
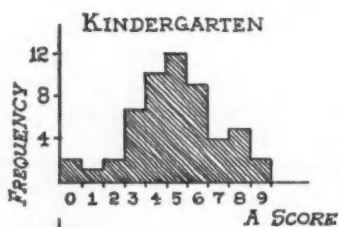


FIGURE 3. A. Relation of point score to grade. B. Relation of A score to grade.



sent study. This factor may be important since a child's ability to deal with numbers might well be determined in part by the amount of formal instruction he has had, and his familiarity with play materials available in school. Clearly, the older children, and those in higher grades, do better than younger children, and better than those in kindergarten (cf. Table II, Figure 3). The two factors are confounded in the present study; to separate them, it would be necessary to compare the abilities of groups of children, matched for age and I.Q., who have been in school for different lengths of time.

Although no difficulty was experienced in classifying an answer as an "A response" or not, it was not so easy to decide whether, over-all, a child was in stage A, B, or C. Moreover, it appears that a child may be in stage A for one type of material and situation, and in stage B or even C for another. This, of course, does not mean that a theory of stages is untenable, but it does suggest that a child may acquire the set of operations necessary for dealing consistently with one type of material and situation without simultaneously being able to deal consistently with all apparently similar situations. The following procedure was therefore adopted: the child's stage was assessed separately for the five different subgroups of the test (Table II). If all the answers in a subgroup indicated judgments based on perceived characteristics, the category assigned was A; if all the answers were operational, the stage assigned was C; if there was a mixture of types of answers or uncertainty about whether responses were fully operational, the category B was used. There were some disagreements between scorers, principally as to whether a child was in stage A or B; however, inter-scorer reliability was over .95.

Figure 4 shows percentages of children in the different stages for three test subgroups. Quite clearly it would be impossible to state a "typical" age for the attainment of concrete operational activity: whereas 80 per cent of children aged 5 years and 10 months showed concrete operations when dealing with provoked correspondence, none could deal in this fashion with unprovoked correspondence, and only 60 per cent with conservation. Another interesting fact is that there appears to be no change in the percentage of children in stage C for provoked correspondence, between the ages of 5;6 and 8;6, but for conservation and unprovoked correspondence there are very marked changes. There are very few children in stage B for provoked correspondence, but stage B represents, over-all, the largest category in unprovoked correspondence. (It is possible that this may have been partly a function of the questions asked—this point would bear further investigation.) The graphs for the other subgroups—seriation and cardinal-ordinal properties—are not so neat as those shown in Figure 4. As one might expect, seriation showed, on the average, more C responses than cardinal-ordinal properties, but

neither showed a very clear age trend. Therefore, although a child can be assigned with a fair degree of assurance to one of the three categories for each test subgroup, there is no consistency of stages within individuals, nor are the age trends similar for the different subgroups. There remains the question of whether ability to operate with serial relations

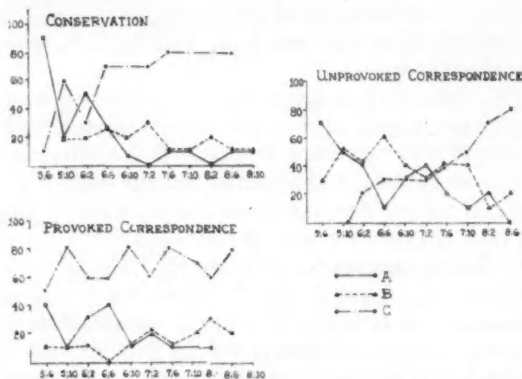


FIGURE 4. Percentages of children at different ages falling into the three stages for three test subgroups; A = global comparison stage, B = intuitive stage, C = operational stage.

and cardinal properties are necessary conditions of being able to deal with numbers (constructs which entail both these properties). If Piaget's thesis is correct, then it should be the case that a child who can deal operationally with cardinal-ordinal properties is also in stage C for unprovoked correspondence and seriation, and, conversely, that a child in stage C for these latter two should have a high probability of being able to deal operationally with cardinal-ordinal properties.

A random sample of 100 was drawn from the original 250 test papers. Of these, 15 showed stage C responses for cardinal-ordinal properties. Table IV shows the conditional probabilities associated with the states of affairs mentioned above. Clearly, Piaget's thesis would require all the conditional probabilities in column I of the table to be unity, or close to it, and those in column II to be 0. A number of the conditional probabilities depart rather markedly from the predicted values. That is, it seems that some children can deal operationally with cardinal-ordinal properties before they can so deal with either classes or series separately, and that ability to deal with classes and series separately does not entail ability to deal with numbers as constructs combining ordinal and cardinal operations.

TABLE IV  
RELATIONS BETWEEN STAGES FOR TEST SUB-  
GROUPS III, IV, AND V

$\frac{P(CIII + CIV)}{(CV)} = 0.56$	$\frac{P(AIII)}{(CV)} = 0.06$
$\frac{p(CIII)}{(CV)} = 0.8$	$\frac{P(AIV)}{(CV)} = 0.00$
$\frac{p(CIV)}{(CV)} = 0.56$	$\frac{P(BIII)}{(CV)} = 0.13$
$\frac{P(CV)}{(CIII + CIV)} = 0.62$	$\frac{P(BIV)}{(CV)} = 0.45$

The measures are conditional probabilities. Letters identify stages, Roman numerals identify test subgroups. Thus  $P(CIII + CIV)/(CV)$  reads: the conditional probability that a subject is in stage C for subgroups III and IV, given that he is in stage C for subgroup V.

#### DISCUSSION

The findings reported above differ from those reported by Estes (1956). According to Estes, children either are able to count, and can deal operationally with situations very similar to those used in the present experiment, or cannot count, and do not respond in terms of numbers. "No stages were found in the development of the number concept—12 of the 14 four-year-olds, 18 of the 20 five-year-olds, and all of the six-year-olds had the correct concept of number. The four children at four and five years who were unsuccessful in counting the chips were also unsuccessful in the blocks, the marbles, and the match sticks. However, they did not show the earlier stages described by Piaget. They just kept taking the chips out to play with them" (Estes, 1956). Estes' report of her procedure is rather inadequate, but apparently the manipulations required of her subjects were very limited. For instance, the children were asked to count ten chips when placed (a) in a row, (b) in a pattern, and (c) in a loose pile. Either children could count all of these, or none of them. However, in Piaget's terms, this does not mean that they have, or have not, the *concept* of number. Children may be able to point to objects, and say numbers in the correct sequence, but the question one should ask is: Do they understand that, having done this once, the number will not change with perceptual transformations? This is the sort of approach which should throw light on the developmental stages (if any) in number concept attainment, but Estes' work is not designed to clarify such questions. The types of response described by Piaget for all three stages were found in the present study, and his "intuitive" stage was quite prevalent, especially for some situations (Figure 4). Perhaps the most striking response in this category, given by several children, was to count the poker

chips when one of the two rows was extended, but still to judge that there were different numbers of chips in the two rows. Thus, counting *per se* is no guarantee that a child grasps what the concept "cardinal number" is, or how it applies to a concrete situation.

The evidence here produced confirms Piaget's contention that young children do not fully understand the concept of number, even though they may be able to count. The sorts of behaviour described by Piaget are typical for children between the ages of five and eight. However, there is inconsistency in the type of response made in different situations, and great variability from child to child, at a particular age level, in the sorts of response made. Intelligence, as well as age, is an important factor (Table III). The stages do not always follow in the sequence Piaget's theory requires (Table IV). Piaget pays scant attention to the part learning may play in the development of number concepts, and points to the fact that children in stages A and B do not understand when one tries to explain to them why their judgments are wrong. But one might suggest that a single explanation cannot be effective since, in Piaget's terminology, the assimilation of the concept explained takes time (Piaget, 1952b, 1954). Churchill (1958) has shown that training with number games over a period of weeks leads to improved understanding of the concept of number. It could be that the variability in types of response from one test subgroup to another is due to incomplete assimilation of a newly learned concept; or, to use Piaget's term again, that the child requires time to accommodate his responses in a novel situation. This could be the case where correct operations have been attained, but not yet integrated into a logical structure, or grouping of operations (Piaget, 1953). Unfortunately, it would be difficult to test this hypothesis, except in an intensive longitudinal study of the development of number concepts, and further progress in this field will necessitate such studies. It seems reasonable to suggest, in the absence of evidence on the point, that the variability in stages, and the absence of strong sequential dependencies predicted from Piaget's theory, is a function of learned responses to particular situations without complete assimilation, or to use a more familiar term, a function of learning without adequate response generalization.

#### SUMMARY

A study of number concepts in 250 children between 5 and 8 years old showed that the three stages of cognitive development described by Piaget as "global," "intuitive," and "concrete operational" occur. There are considerable variations in type of response given at any age level, and the type of response may vary from one test situation to another for a child. There are age trends; older children tending to give more "operational" judgments, but these trends differ for different test situa-

tions. Intelligence, as measured on a standard group I.Q. test, is also a factor in number concept attainment.

The findings do not yield unequivocal support for Piaget's theory of cognitive development. Possible reasons for this are discussed, and a strategy for further research suggested.

#### REFERENCES

- CHURCHILL, E. M. The number concepts of the young child. *Researches and Studies, University of Leeds Inst. of Educ.*, 1958, 17, 34-49.
- DODWELL, P. C. The evolution of number concepts in the child. *Mathematics Teaching*, 1957, 5, 5-11.
- ESTES, B. W. Some mathematical and logical concepts in children. *J. Genet. Psychol.*, 1956, 88, 219-222.
- HAZLITT, V. Children's Thinking. *Brit. J. Psychol.*, 1930, 20, 354-361.
- MCCARTHY, D. Language development of the preschool child. *Inst. Child Welf. Monogr. No. 4*. Minneapolis: Univ. of Minnesota Press, 1930.
- PIAGET, J. *The language and thought of the child*. New York: Harcourt, 1926.
- PIAGET, J. *The child's conception of the world*. New York: Harcourt, 1929.
- PIAGET, J. *The psychology of intelligence*. London: Routledge and Kegan Paul, 1950.
- PIAGET, J. *The child's conception of number*. London: Routledge and Kegan Paul, 1952.
- PIAGET, J. *The origins of intelligence in children*. New York: International Univer. Press, 1952.
- PIAGET, J. *Logic and psychology*. Manchester: Manchester Univer. Press, 1953.
- PIAGET, J. *The construction of reality in the child*. New York: Basic Books, 1954.
- RUSSELL, B. *Introduction to mathematical philosophy*. London: Allen & Unwin, 1919.

## THE EFFECTS OF NON-REINFORCEMENT ON RESPONSE STRENGTH AS A FUNCTION OF NUMBER OF PREVIOUS REINFORCEMENTS<sup>1</sup>

RONALD K. PENNEY<sup>2</sup>

*State University of Iowa*

THE OMISSION of a customary reward has been observed by many investigators to lead to an increase in the strength of a previously rewarded response (Miller, 1936; Rohrer, 1949; Sheffield, 1950, 1954). These investigators have all entertained some notion of the motivational effects of non-reward (the frustration hypothesis) in interpreting their observations. Although these studies support the frustration hypothesis, they are not crucial since alternative hypotheses of this increased vigour are readily available in terms of associative factors. Brown and Farber (1951) have reviewed a number of "non-motivational" interpretations of frustration. For example, the subject may have learned in the experimental situation, or in similar situations in his previous history, to make more effortful responses in the presence of non-reinforcement cues.

In an effort to restrict the number of alternative interpretations, Amsel and his associates (Amsel & Hancock, 1957; Amsel & Roussel, 1952; Roussel, 1952) utilized a two-runway situation in which the criterion response was spatially different from the non-reinforced instrumental response. Amsel and Roussel trained rats to run down an alley into a goal box and then into a second alley to another goal box. Performance in the second alley, when reward was omitted from the first goal box, was superior to performance on test trials when reward was received at this goal box. The facilitative effect that this non-reinforcement had on the subsequent response was termed the frustration effect. Roussel (1952) also used the two-runway situation and studied the development of the frustration effect under conditions of partial reinforcement. Roussel found that the frustration effect was not immediate but only developed after a certain number of reinforced trials. Wagner (1959) followed up the Tulane studies and confirmed the developing nature of the frustration effect when the first runway response was partially reinforced.

Both the Roussel and Wagner studies imply that the frustration effect is a function of the number of reinforcements prior to non-reinforcement.

<sup>1</sup>Part of a dissertation submitted to the Graduate College of the State University of Iowa in partial fulfillment of the requirements for the Ph.D. degree. The author is indebted to Alfred Castaneda for his advice and assistance throughout the course of the investigation.

<sup>2</sup>At present at McMaster University, Hamilton, Ontario.

Two studies (Holton, 1956; Marzocco, 1950) have systematically investigated the effect of non-reinforcement when the number of continuous reinforcements was varied prior to non-reinforcement. Holton, using children, found the increased vigour of responding following non-reinforcement to be dependent on the number of previous reinforcements. Marzocco, using rats, did not find a relation between the vigour of responding following non-reinforcement and the number of prior reinforcements. Neither study utilized a criterion response that was spatially different from the non-reinforced instrumental response.

The present study was designed to determine whether the frustration effect is a function of the number of continuous reinforcements prior to non-reinforcement when children are used as subjects. A modification of Amsel's technique for studying the frustration effect was adopted for use with children.

### METHOD

#### *Subjects*

The Ss were 88 kindergarten children obtained from two elementary schools in Washington, Iowa. Twenty-two Ss were randomly assigned to each of the four treatment groups.

#### *Apparatus*

Figure 1 is a schema of the apparatus. The face of the apparatus consists of two levers ( $R_1$  and  $R_2$  in Figure 1), two stimulus lights ( $S_1$  and  $S_2$ ) and a goal box ( $GB_2$ ). In addition, a clear piece of plastic tubing was attached to the side of the apparatus and extended down past the base of the apparatus. The entire apparatus was painted flat black with the exception of the two levers, which were flat grey.

$R_1$  moved to the left of S through an excursion of 5 in.  $R_2$  moved towards S through an excursion of 11 in. At the completion of  $R_2$ 's excursion, a marble was released at  $GB_2$ . At the completion of  $R_1$ 's excursion, a marble dropped into the clear plastic tubing on the rewarded trials. A door chime sounded at the same time as the marble was released into the plastic tube.

The two stimulus lights,  $S_1$  and  $S_2$ , served to elicit responses to  $R_1$  and  $R_2$  respectively, and were turned off at the completion of these responses.

As shown in Figure 1, a hand pattern was located on the lower right corner of the face of the apparatus. The S was required to place his hand on this pattern after he had manipulated  $R_1$ . The onset of  $S_2$  served as a signal for S to remove his hand from the hand pattern and pull down on  $R_2$ .

Additional material included assorted candy of the chocolate bar and gum drop variety. At the beginning of each experimental session, one chocolate bar and one gum drop were placed in the clear plastic container immediately above the plastic tubing.

#### *Procedure*

The S was brought into the experimental room and the apparatus was introduced as a marble game. He was instructed that if he filled the clear plastic tube with marbles, he could have the two pieces of candy in the container immediately above the plastic tube.



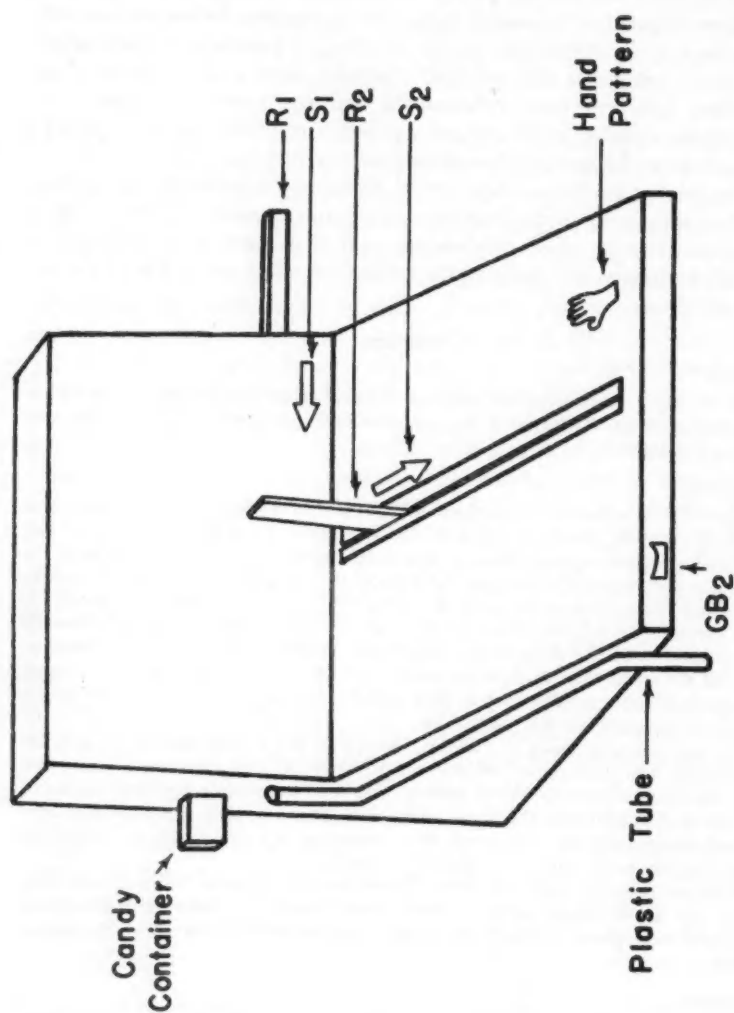


FIGURE 1. Schematic drawing of the apparatus.

The experimental period consisted of a training phase and a testing phase. There were three phases of training.

#### *Training Phase I*

Each S was given four training trials in Phase I. The details of procedure on an individual trial are as follows: (a) the first stimulus light ( $S_1$ ) was illuminated; (b) S pushed  $R_1$  and at the completion of  $R_1$ 's excursion a marble dropped into the plastic tube; (c) S placed his hand on the hand pattern; (d) the second stimulus light ( $S_2$ ) was illuminated at the end of 10 sec.; (e) S removed his hand from the hand pattern and pulled  $R_2$  through its excursion; (f) a marble was released at the goal box ( $GB_2$ ) when  $R_2$  completed its excursion; (g) S placed this marble in the plastic tube. All Ss received four reinforced trials to  $R_1$  and four reinforced trials to  $R_2$  in the sequence indicated above, during Phase I.

#### *Training Phase II*

During Phase II of training, differential treatment was administered in that the Ss were divided into two groups. The Hi-Habit Group received 10 trials on  $R_1$  alone (i.e., not followed by trials on  $R_2$ ). The Lo-Habit Group received one trial on  $R_1$  and similarly did not receive any trials on  $R_2$ .

#### *Training Phase III*

In order to re-establish the  $R_1 - R_2$  sequence of responding, two trials were administered following the same details of procedure as outlined in Phase I.

#### *Testing Phase*

During the testing phase, the two groups were further divided into two additional groups—an experimental group and a control group. The experimental group received 18 trials, where a trial consisted of an  $R_1$  response followed by an  $R_2$  response, as in the first phase of training. For the experimental group the details of an individual trial were the same as those outlined in Phase I of training, except that on 12 of these trials, a marble did not drop into the plastic tube when S pushed  $R_1$  through its excursion. The trials on which the marble failed to appear were designated the non-reinforced trials. Test trials 1, 2, 3, 4, 6, 8, 9, 10, 12, 15, 16, and 18 represented the non-reinforced test trials. Thus the experimental groups of the Hi-Habit and Lo-Habit conditions received 12 non-reinforced trials and six reinforced trials on  $R_1$  during the testing phase. On the other hand, the control groups of the Hi-Habit and Lo-Habit conditions were always reinforced for an  $R_1$  response and an  $R_2$  response. The control groups, therefore, received 18 reinforced trials to both manipulanda.

When non-reinforcement was first introduced, the marbles in the plastic tube had accumulated to the same location in the tube. This was accomplished by affixing a piece of plastic tubing to the bottom of the plastic tube at the beginning of the experimental session for the Hi-Habit groups. This extension contained the nine additional marbles that the Hi-Habit groups received prior to the first non-reinforcement. After the first non-reinforcement, however, the control groups were closer to filling the plastic tube than were the experimental groups as the control groups received two marbles on every trial. The measurements of performance were made in relation to  $R_2$ . A movement time measure involved the time it took S to pull  $R_2$  through its excursion.

## RESULTS

In analysing the  $R_2$  movement time measure, the reciprocal of each measure multiplied by ten was obtained for each subject. These scores were referred to as movement speed. Figure 2 presents the movement

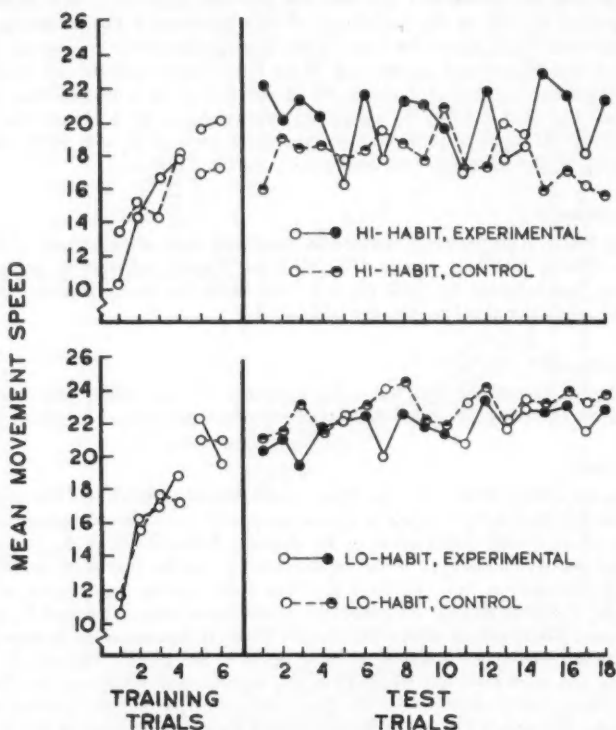


FIGURE 2. Mean  $R_2$  movement speeds plotted over the reinforced (open circles), nonreinforced (solid circles) and corresponding (semi-solid circles) trials for each condition.

speed curves for the Hi-Habit experimental and control groups and the Lo-Habit experimental and control groups. Considering these curves, it is evident that the Hi-Habit experimental group shows no overlap between the reinforced and non-reinforced test trials while all the other groups exhibit considerable overlap. Furthermore, greater movement speed is associated with the non-reinforced trials rather than the reinforced trials for the Hi-Habit experimental group. Table I presents the

mean movement speeds over the last two training trials, all the reinforced test trials and all the non-reinforced trials.

In order to determine whether the groups differed in movement speed prior to differential treatment, the first six training trials were blocked in trials of one and an analysis of variance (Lindquist, 1953) was performed. None of the sources of variation was significant except the effect of trials. Similarly, an analysis of variance was performed over the mean of the last two training trials for all four groups and again no reliable difference was found.

TABLE I

MEANS AND SD'S OF MOVEMENT SPEEDS OVER THE LAST TWO TRAINING TRIALS, THE REINFORCED TEST TRIALS, AND THE NON-REINFORCED OR CORRESPONDING TEST TRIALS

Group	Last two training trials		Reinforced test trials		Non-reinforced or corresponding trials	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Hi-Habit Experimental	19.92	7.01	17.59	6.57	21.34	7.68
Hi-Habit Control	17.15	5.55	18.39	6.50	17.84	5.53
Lo-Habit Experimental	20.94	6.41	21.53	6.54	21.69	6.37
Lo-Habit Control	20.85	7.98	23.11	8.81	22.83	7.58

To determine the effect of non-reinforcement on movement speed during testing, an analysis of variance was conducted over the test trials for all four groups. For the experimental groups this analysis involved a comparison of the differences in mean movement speeds between the reinforced and non-reinforced test trials. In the case of the control groups, however, the analysis compared differences in mean movement speeds between the reinforced test trials and the reinforced test trials corresponding to the non-reinforced trials of the experimental group. Table II summarizes this analysis and, as indicated, a reliable triple interaction was found suggesting that the non-reinforcement effect depended on the level of habit. Accordingly, the above comparisons were repeated for each level of habit separately. Under the Lo-Habit condition, all *F*s were less than one. However, under the Hi-Habit condition a reliable ( $F = 20.57$ ,  $df = 1, 42$ ,  $p < .001$ ) test trials by experimental versus control interaction was found. This interaction indicated that the difference in mean movement speeds between the reinforced and the non-reinforced test trials was greater for the experimental group alone. Comparing the reinforced with the non-reinforced trials a reliable difference ( $t = 7.08$ ,  $df = 21$ ,  $p < .001$ ) was found for the experimental group with the faster mean movement speed being associated with the non-reinforced test trials. A similar *t*-test was computed for the control

TABLE II

SUMMARY OF THE ANALYSIS OF VARIANCE OF MEAN MOVEMENT SPEEDS OVER ALL THE REINFORCED TEST TRIALS AND ALL THE NON-REINFORCED OR CORRESPONDING TEST TRIALS FOR ALL FOUR GROUPS

Source	df	ms	F	p
Between subjects	87			
Habit	1	547.07	5.52	
E vs. C	1	0.01	—	
Habit $\times$ E vs. C	1	78.31	—	
Error (b)	84	99.12		
Within subjects	88			
Test trials	1	27.23	6.90	<0.01
Test trials $\times$ Habit	1	28.75	7.26	<0.01
Test trials $\times$ E vs. C	1	64.46	16.28	<0.001
Test trials $\times$ Habit $\times$ E vs. C	1	39.13	9.88	<0.005
Error (w)	84	3.96		
TOTAL	175			

group and no reliable difference in mean movement speeds was found between the test trials corresponding to the reinforced and the non-reinforced test trials of the experimental group.

Inspection of the Hi-Habit experimental and control group curves of Figure 2 suggested that the effect of non-reinforcement depended on the stage of testing. Accordingly, an analysis of variance was performed on mean movement speeds of the twelve non-reinforced test trials of the Hi-Habit experimental group and the corresponding trials of the Hi-Habit control group. The test trials by experimental versus control interaction was reliable ( $F = 2.30$ ,  $df = 11$ ,  $462$ ,  $p < .01$ ) and suggested that the difference between the experimental and control groups was a function of the particular trial. A series of  $t$ -tests computed for each trial yielded reliable differences between the experimental and control groups on trials 1, 10, 11 and 12 of the non-reinforced test trials, that is, on the first and last three non-reinforced trials. In each case the experimental group's performance was superior to that of the control group. On the other hand, when the *reinforced* test trials were compared with the corresponding trials of the control group no reliable differences in mean movement speed were found.

#### DISCUSSION

The main finding of the study is that non-reinforcement of a response,  $R_1$ , increases the speed of a subsequent (10 sec. later) response,  $R_2$ , when both  $R_1$  and  $R_2$  have been reinforced by the same reward (marble). Further, the increment in speed of  $R_2$  was found to be a function of

the number of continuous reinforcements prior to the introduction of non-reinforcement. It should be noted, however, that the increment in  $R_2$  speed was found to be reliable only on the first and the last three non-reinforced trials. Under the experimental conditions of the present study, the effect of non-reinforcement is not stable but relatively transient.

Amsel (1952) suggested that the non-reinforcement of a previously reinforced response (frustration) results in an increase in generalized drive strength. According to Hull (1943) an increment in generalized drive strength would increase the strength of some ongoing response provided the drive conditions did not lead to a response which was antagonistic to the ongoing response. Applying Amsel's motivational interpretation to the present study, the non-reinforcement of  $R_1$  introduced frustration into the motivational complex and strengthened the subsequent response  $R_2$ . However, frustration appeared to be dependent on the number of reinforced  $R_1$  trials prior to non-reinforcement. Some expectation or an anticipatory goal reaction may be a necessary condition for frustration and may be dependent on the number of reinforced trials prior to non-reinforcement.

The frustration interpretation meets some difficulties, although not insurmountable, in "explaining" a number of other aspects of the results. First, if a given number of reinforcements prior to non-reinforcement is essential to frustration, why does the Lo-Habit experimental group not show a frustration trend by the end of the test trials? This aspect is not inconsistent with a frustration interpretation if partial reinforcement does not lead to the same motivational increment as continuous reinforcement. That is, although the Lo-Habit experimental group has received thirteen reinforced  $R_1$  trials by the end of testing, these reinforcements have been interspersed with ten non-reinforced  $R_1$  trials. It is quite likely that for humans the partial condition produces less generalized drive strength than a series of continuous rewards prior to non-reward. This hypothesis is testable and would involve comparing a group that received 50 per cent  $R_1$  reward from the beginning of training with a group that received 100 per cent  $R_1$  reward prior to the first non-reward.

The second difficulty that the frustration hypothesis encounters concerns a possible sampling error. Evidence for sampling error may be found in the performance trend of the Hi-Habit experimental and control groups over the last two training trials. The Hi-Habit control group exhibited slower (not significantly) mean  $R_2$  speeds relative to the Hi-Habit experimental groups. However, even assuming that the between-groups effect was due to sampling error, the within-groups effect (reinforced *versus* non-reinforced trials) must be accounted for and is con-

sistent with a frustration interpretation. That is, the Hi-Habit experimental group exhibited faster  $R_2$  speeds over the non-reinforced trials relative to the reinforced trials.

One other aspect of the data should be mentioned. Inspection of Figure 2 reveals that the  $R_2$  speed of the Hi-Habit groups was slower than the  $R_2$  speed of the Lo-Habit groups over the test trials. The depressed performance of the Hi-Habit groups may be due to fatigue resulting from the additional nine trials they received. One argument against such a factor is that according to this fatigue hypothesis the Lo-Habit groups should have shown a comparable depression approximately midway through the test trials. No such depression was found. However, in support of the fatigue hypothesis, it should be noted that almost twice as much pressure was required to push  $R_1$  through its excursion as to push  $R_2$ . The Hi-Habit groups made ten  $R_1$  responses in succession while the  $R_1$  responses of the Lo-Habit groups were interspersed with  $R_2$  responses and rest periods. Under these conditions it is conceivable that the fatigue factor may have been less for the Lo-Habit groups. At any rate, since the frustration effect involves a comparison between the experimental and control groups at each level of habit (a within-habit effect) this factor is not crucial to the frustration interpretation.

#### SUMMARY

The present investigation studied the effects of non-reinforcement on response speed as a function of the number of reinforcements prior to non-reinforcement. Eighty-eight kindergarten children were trained to manipulate a lever ( $R_1$ ) and receive a marble and then manipulate a second lever ( $R_2$ ) and receive another marble. Differential training on  $R_1$  was interspersed amongst this  $R_1$ - $R_2$  reinforcement sequence. The Hi-Habit Group received nine more trials than did the Lo-Habit Group. During the test series, the Hi-Habit and Lo-Habit groups were randomly divided into two groups—an experimental group and a control group and responded in an  $R_1$ - $R_2$  sequence. However, for the experimental groups a marble was omitted following an  $R_1$  response on 12 of the 18 test trials. The control groups continued to receive a marble after each  $R_1$  response. The time it took to begin to manipulate  $R_2$  after a signal was presented (starting time) and the time it took to manipulate  $R_2$  were recorded on each trial.

The Hi-Habit experimental group showed reliably faster mean movement speeds over the non-reinforced trials relative to the reinforced trials when the differences in mean movement speeds were compared for all groups. In addition, the mean movement speeds of the Hi-Habit experimental and control groups were compared over the non-reinforced trials and corresponding trials. The experimental group was found to perform significantly faster over the first and last three non-reinforced trials. The Lo-Habit groups did not show any non-reinforcement effect over the test series.



## REFERENCES

- AMSEL, A. The role of frustrative nonreward in noncontinuous reward situations. *Psychol. Bull.*, 1958, **55**, 102-119.
- AMSEL, A., & HANCOCK, W. Motivational properties of frustration: III. Relation of frustration effect to antedating goal factors. *J. exp. Psychol.*, 1957, **53**, 126-131.
- AMSEL, A., & ROUSSEL, J. Motivational properties of frustration: I. Effect of a running response on the addition of frustration to the motivational complex. *J. exp. Psychol.*, 1952, **43**, 363-368.
- BROWN, J. S., & FARBER, I. E. Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, **48**, 465-495.
- HOLTON, R. B. Variables affecting the change in instrumental response magnitude after reward cessation. Unpublished Ph.D. dissertation, State University of Iowa, 1956.
- HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
- LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.
- MARZOCCO, F. N. Frustration effect as a function of drive-level, habit strength and distribution of trials during extinction. Unpublished Ph.D. dissertation, State University of Iowa, 1950.
- MILLER, N. E., & STEVENSON, S. S. Agitated behavior in rats using experimental extinction and a curve of spontaneous recovery. *J. comp. Psychol.*, 1936, **21**, 205-231.
- ROHRER, J. H. A motivational state resulting from nonreward. *J. comp. physiol. Psychol.*, 1949, **42**, 476-485.
- ROUSSEL, J. Frustration effect as a function of repeated nonreinforcements and as a function of consistency of reinforcement prior to the introduction of non-reinforcement. Unpublished master's thesis. Tulane University, 1952.
- SHEFFIELD, V. F. Resistance to extinction as a function of the spacing of extinction trials. *J. exp. Psychol.*, 1950, **40**, 305-313.
- SHEFFIELD, F. D., & CAMPBELL, B.A. The role of experience in the spontaneous activity of hungry rats. *J. comp. physiol. Psychol.*, 1954, **47**, 97-100.
- WAGNER, A. R. The role of reinforcement and nonreinforcement in an "apparent frustration effect." *J. exp. Psychol.*, 1959, **57**, 130-136.

## DISTRIBUTION VARIABLES IN SIMPLE DISCRIMINATION LEARNING IN RATS<sup>1</sup>

M. R. D'AMATO  
*New York University*

SEVERAL RECENT STUDIES (for example, Riopelle & Churukian, 1958; Sarason, Sarason, Miller, & Mahmoud, 1956; Thompson & Pennington, 1957) have reported inconsistent results concerning the role of the inter-trial interval (ITI) in discrimination learning, indicating the operation of unknown relevant variables. In many studies, for example, the subject is handled from trial to trial, in removal from the goal box, etc.; and it is quite possible that recovery from the inter-trial handling might be a factor in the facilitative effects sometimes shown by spacing of trials. Another possible relevant variable is the stimulus situation to which the subject is exposed during the ITI. If these stimuli are similar to those operating within the experimental situation, reactions conditioned to them during the ITI could generalize to the latter, influencing the course of learning.

In the present study the role of the ITI is studied in a situation where (a) handling during the ITI is eliminated, (b) the visual stimuli acting during the ITI are varied between two extremes, and (c) a third variable, trials per day, also comes under consideration.

### METHOD

#### *Subjects*

The Ss were 72 experimentally naïve albino rats (42 females and 30 males) obtained from litters bred in our laboratory. Their average age at the start of the study was about 80 days.

#### *Apparatus*

The apparatus consisted of four units of an automatic learning-discrimination apparatus described in a previous report (D'Amato & Jagoda, 1960). Essentially, this consists of an enclosed symmetrical Y maze (arms set 120 degrees apart) with remote programming and recording equipment. The use of a symmetrical Y maze permits the same arm to be used as a "goal" and a "starting" arm, allowing for a sequence of automatically controlled trials. For example, in a brightness discrimination problem the doors open and S chooses between a bright and a dimly illuminated arm. The arm of his choice becomes, on the next trial, the "starting" arm, the discriminanda appearing in the remaining two arms. Each trial runs off in this manner in accordance with a pre-established programme until the day's trials are completed.

<sup>1</sup>This research was supported by research grant M-2051 from the National Institutes of Health, Public Health Service, Department of Health, Education, and Welfare, Bethesda, Maryland. The assistance of H. Jagoda is gratefully acknowledged.

### Procedure

**Pretraining.** Four days before the beginning of discrimination training, Ss were placed on a 23 hr. water deprivation training regimen, food being available *ad lib*. Three days later each S was placed in an arm of the Y maze with lights out and doors closed and permitted to drink five dipperfuls of water; three dipperfuls were allowed on the next day, the first day of training.

**Discrimination training.** The 72 Ss were equally divided among the 12 cells resulting from the three experimental variables: inter-trial interval; trials per day; and stimulus conditions during the ITI. There were two values of the inter-trial interval, 30 sec. (range 24-36 sec.) and 102 sec. (range 96-108 sec.). The number of trials per day was 10, 20, or 40. One of two conditions prevailed during the ITI: all lights in the Y maze were extinguished ("lights out"—LO), or the discriminanda remained on throughout the ITI ("lights"—L).

For all Ss,  $S_+$  was the brightly illuminated arm (10.0 ft.-c.), and  $S_-$  was the dimly lighted arm (about .01 ft.-c.), measured under our standard conditions. All Ss were run until the criterion of 18 correct responses out of 20 successive trials was met.

On a correct choice the sequence of events was as follows. The S was permitted 1.5 sec. drinking (reward interval), timed from S's first lap of water, after which the dipper retracted out of sight. The discriminandum,  $S_+$  in this case, remained on in the arm for an additional 8 sec. ("stimulus" interval). The inter-trial interval, either 30 or 102 sec., then followed. All lights in the Y maze were extinguished during the ITI for the LO groups; for the L groups, however, they remained on. Finally, a flashing light during a 5 sec. ("starting") interval signalled the beginning of the next trial. On an incorrect choice the sequence of events was identical, except that the dipper was retracted during the initial 1.5 sec. period.

The estimated average duration of water deprivation at the start of a day's trials was 22.5 hr. During the experiment the temperature in the laboratory ranged between 72-78° F. and the relative humidity varied between 50-65 per cent.

### RESULTS AND DISCUSSION

The mean numbers of trials to criterion for the various groups are presented in Table I. After establishing that the requirement of homogeneity of variances was met, an analysis of variance was applied to the criterion scores, yielding two significant *F*s. An *F* of 12.29 ( $p < .001$ ) was obtained for L *versus* LO variable, and an *F* of 4.74 ( $p < .05$ ) for the L  $\times$  ITI interaction. The nature of the interaction is apparent from Table I. Although the LO condition is superior to the L condition for both ITI's, the superiority is vastly greater for the shorter ITI. The interaction is also indicative of the fact that the ITI is a relevant variable only for the L condition, that is, the facilitative effect of the longer ITI vanishes when the discriminanda are removed during the ITI. A similar analysis was performed on the errors to criterion scores with precisely the same results. The *F*s in this case were 16.88 and 5.68, respectively.

While the  $F$  for the trials per day variable did not approach significance, the results presented in Table I offer some reassurance for our standardized training procedure of 20 trials per day.

TABLE I  
MEAN NUMBER OF TRIALS TO CRITERION

ITI	Trials per day			Stimuli during ITI	
	10	20	40	L	LO
30 sec.	88.8	78.5	82.0	102.1	64.1
102 sec.	76.9	64.0	76.5	75.6	66.7

The data clearly indicate that, within the limits of the present experiment, the ITI ceases to be a variable of importance when all relevant stimuli are removed during the interval. This result provides a clue as to when the ITI is likely to be of consequence in other experiments, that is, when the ITI contains stimuli that are relevant to the training situation, such as discriminanda, extra-maze cues, and so forth. It is interesting that Riopelle and Churukian (1958), using ITI's of the same order as those of the present study, found this variable to be without effect on discrimination learning in monkeys. In their study the subject spent the ITI, unhandled, in his home cage with an opaque screen barring all visual cues from the experimental situation.

One possible interpretation of the preceding result is that the presence of relevant stimuli during the ITI allows for the conditioning of responses that subsequently interfere with correct responding. For example, competing ("frustration") responses could, in the L condition, become conditioned to  $S_+$  during the ITI, while the subject awaits the starting of the next trial (cf. Adelman & Maatsch, 1955). Such responses would in some measure interfere with approach to  $S_+$ . Since the L condition is more detrimental for the shorter ITI, it is presumed that these competing responses reach a maximum relatively early in the ITI and, therefore, after a short ITI, are available on the following trial at considerably greater intensity than after a long ITI.

Finally, the fact that the LO condition is superior to the L condition for both ITI's implies that the subject learns little about the discrimination problem through mere exposure to the discriminanda. Apparently, preservation of the initial responses (based on the reinforcement contingencies) to the discriminanda by removal of the latter is a factor of far greater importance.

## SUMMARY

Seventy-two rats were trained on a brightness discrimination in an automatic Y maze under two inter-trial interval (ITI) durations (30 versus 102 sec.) and with 10, 20, or 40 trials per day. The third variable in the  $2 \times 3 \times 2$  design was the presence or absence of the discriminanda during the ITI. Rate of acquisition of the discrimination (a) was found to be significantly faster with elimination of the discriminanda during the ITI, the more marked effect occurring for the shorter ITI, and (b) was faster with the longer ITI when the discriminanda were present during the ITI but was independent of the ITI when the discriminanda were absent.

## REFERENCES

- ADELMAN, H. M., & MAATSCH, J. L. Resistance to extinction as a function of the type of response elicited by frustration. *J. exp. Psychol.*, 1955, 50, 61-65.
- D'AMATO, M. R., & JAGODA, H. Effects of extinction trials on discrimination reversal. *J. exp. Psychol.*, 1960, 59, 254-260.
- RIOPELLE, A. J., & CHURUKIAN, G. A. The effect of varying the intertrial interval in discrimination learning by normal and brain-operated monkeys. *J. comp. physiol. Psychol.*, 1958, 51, 119-125.
- SARASON, I. G., SARASON, B. R., MILLER, M., & MAHMOUD, P. The role of the intertrial interval in discrimination and reversal learning. *J. comp. physiol. Psychol.*, 1956, 49, 77-79.
- THOMPSON, R., & PENNINGTON, D. F. Memory decrement produced by ECS as a function of distribution in original learning. *J. comp. physiol. Psychol.*, 1957, 50, 401-404.

## SOME DATA RELATING TO THE POSSIBILITY OF USING A SHORTER FORM OF THE HEBB-WILLIAMS TEST

J. J. LAVERY, AND D. BÉLANGER

*Université de Montréal*

IN A RECENT ARTICLE, Das and Broadhurst (1959) have presented data demonstrating order effects on the difficulty level of most of the Hebb-Williams test items. They stress the need for further analysis of this order variable and also suggest that two items be changed. It is the purpose of the present paper (a) to submit results indicating that a shorter form of the test, while more practical, would be just as reliable as the long form, and (b) to propose that the possibility of reducing the length of this test be fully investigated before determining the difficulty level and best possible order of administration.

Although Rabinovitch (1949) rightly remarks that the Hebb-Williams test is considerably shorter than "traditional mazes and problem boxes," the administration of this test in its present form still requires a considerable amount of time. To test one group of 36 rats in one session would, for example, take approximately 5½ to 6 hours per session or 12 hours per day; on the other hand, if the rats were divided in three groups of 12 to be tested successively, a period of a month and a half would elapse before all animals would be processed.

This last procedure was adopted in an experiment (Lavery, 1958), the purpose of which was to test the hypothesis that rats with a high emotionality index on the Hall open-field test (Broadhurst, 1957; Hall, 1934) and a high intelligence score on the Hebb-Williams maze would learn to avoid a noxious stimulus faster than rats scoring low on one or both of these tests. The results indicated the existence of a positive correlation between the scores on the intelligence test and the performance in the avoidance situation.

Since learning to avoid a noxious stimulus can be a difficult and long task for rats, the delay caused by the intelligence testing procedure was a serious handicap in this experiment. It was therefore decided to analyse the results of this testing and to investigate the possibility of using a shorter test in the future. There are two possible ways of reducing the length of such a test. First, some problems may be eliminated and the adaptation series shortened. A second method would be to reduce the number of trials per problem in order to allow more rats to be tested in one day. Shortening the adaptation period could very well change the fundamental nature of the test, and the elimination of problems would probably seriously upset the difficulty level of subsequent problems.

Therefore, it was judged preferable to investigate the effect of reducing the number of trials instead of applying this first solution.

A graphical analysis of the scores of these 36 rats on the test immediately suggested that, on most problems, the last four trials did not affect, to any great extent, the distribution of the scores. Therefore, a Pearson  $r$  was calculated between every cumulated score after each trial, and the total score. Furthermore, since 20 of the above-mentioned rats were tested in an avoidance-learning situation, the number of trials required by these rats before obtaining ten avoidance responses out of ten trials in one session was used as an external measure with which to correlate these scores. This measure is not really an external criterion as it would be used in the validation of a test. In the present analysis the validity of the test is assumed and therefore a shorter form which correlates highly with it should also be valid. The correlations between each of the shorter forms and the external measure represent an attempt at verifying this assumption.

The most relevant of these correlations are shown in Table I. These results show that the cumulated score of the first four trials has a high

TABLE I

CORRELATIONS OF PARTIAL SCORES ON THE HEBB-WILLIAMS TEST WITH TOTAL SCORE AND EXTERNAL CRITERION

	<i>N</i>	Total score	First 2 trials	First 3 trials	First 4 trials	Last 4 trials	Trials 2, 3, 4
Total score ( $r$ )	36	—	0.74	0.85	0.91	0.74	0.89
External criterion ( $Rho$ )	20	0.41*	0.39*	0.39*	0.40*	0.37	0.39*

\*Significant at the 0.05 level.

correlation with the total score (.91). On the other hand, the correlation between the last four trials and the total score is much lower (.74), a fact which seems to indicate that the last trials do not contribute as strongly to the total score. Furthermore, the correlation between the first four trials and the external measure (.40) is about the same as that between the total score and this measure (.41). This correlation is somewhat lower for the last four trials (.37). The effect of Hebb and Williams' (1946) original suggestion of eliminating the first trial from the scoring was also investigated. It is evident from Table I that the combined score of trials 2, 3, and 4 yield the same results as the first four trials taken together.

This analysis seems to indicate that the administration of only half of the test would have been just as effective. Even then only trials 2, 3 and 4 need have been scored. It could be objected, however, that such a procedure would reduce the reliability of the score. To verify this point, two



indices of internal consistency were taken for each combination of scores. A Spearman-Brown reliability coefficient was calculated between odd- and even- numbered problems and a Kendall coefficient of concordance ( $W$ ) between all trials in each combination. These results are presented in Table II. They would seem to imply that the shorter forms are

TABLE II  
INTERNAL CONSISTENCY INDICES OF TOTAL AND PARTIAL SCORES ON HEBB-WILLIAMS TEST

	<i>N</i>	Total score	First 2 trials	First 3 trials	First 4 trials	Last 4 trials	Trials 2, 3, 4
Between problems (S-B)	36	0.68	0.16	0.45	0.65	0.65	0.57
Between trials (W)	36	0.40	0.10	0.55	0.53	0.48	0.60

just as consistent from problem to problem and, perhaps even more stable from trial to trial, than the complete test.

Although Rabinovitch (1949, 1951) does not explicitly state in his standardization reports why he decided to give five more trials than Hebb and Williams did in the original procedure, we can surmise that he wanted to increase the reliability of the test. However, since the last four trials have a relatively low correlation ( $r = .39$ ) with the first four and since the addition of these last four trials reduces the inter-trial consistency from .53 to .40, it is not very evident that these extra trials do improve the reliability. The only explanation we can think of is that the last four trials reduce the variability of the total scores. Should this be the case, we might expect the shorter form to yield a better discrimination. It might seem strange that two measures which correlate with each other and which are both correlated to a third do not, when combined, yield a higher correlation with the third. This is not so surprising when it is remembered that these measures are both obtained from the same rats running through the same problems. The only difference between the two is that the second measure is the score obtained after another four trials. In other words, both measure the same thing. This result only emphasizes the main purpose of this note. If the last four trials add no discriminative power to the test they should be dropped; this would make the test more practical, providing, of course, that the reliability is not affected. In any event, the present data justify further experimentation on the possibility of reducing the length of the Hebb-Williams test. The effect of a shorter test on the reliability and on the difficulty level of each problem could be assessed by repeating Rabinovitch's procedure, using alternate sets of problems with four trials instead of eight.

It should be emphasized that no claim is made to the necessity of changing the scoring key of the present test. Such a decision should be based on the use of an external criterion of a more general nature, as well as on a cross-validation study. It is merely suggested that a shorter form of the test may be just as reliable and more practical than the present one. How this shorter test would influence the difficulty level of each item and the reliability of the entire test can only be answered by further experimentation.

#### Résumé

Pour des raisons pratiques, les auteurs ont étudié la possibilité d'abréger le test Hebb-Williams. Les résultats, semblent indiquer que, en utilisant quatre essais au lieu de huit à chaque problème, le test donnerait un indice d'intelligence aussi valide et probablement aussi fidèle. Ces résultats, cependant, ne peuvent que suggérer une étude subséquente dans laquelle la validité réelle du test dans sa nouvelle forme, de même que la difficulté actuelle de chaque problème, pourraient être évaluées.

#### REFERENCES

- BROADHURST, P. L. Determinants of emotionality in the rat. *Brit. J. Psychol.* 1957, 48, 1-12.
- DAS, G. & BROADHURST, P. L. A note on the Hebb-Williams test of intelligence in the rat. *Canad. J. Psychol.*, 1959, 13, 72-75.
- HALL, C. S. Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. *J. comp Psychol.*, 1934, 18, 385-403.
- HEBB, D. O., & WILLIAMS, K. A. A method of rating animal intelligence, *J. gen. Psychol.*, 1946, 34, 59-65.
- LAVERY, J. J. L'influence des facteurs anxiété et intelligence sur la performance du rat dans une situation échappement-évitement. Thèse de licence (dactylographiée), Université de Montréal, 1958.
- RABINOVITCH, M. S. Standardization of a closed field intelligence test for rats. M.Sc. thesis, McGill University, 1949.
- RABINOVITCH, M. S., & ROSVOLD, H. E. A closed field intelligence test for rats. *Canad. J. Psychol.*, 1951, 5, 122-128.

## BOOK REVIEWS

*Electronic Instrumentation for the Behavioral Sciences.* By CLINTON C. BROWN and RAYFORD T. SAUCER. American Lecture Series. Springfield, Ill.: Thomas [Toronto: Ryerson Press]. Pp. xiii, 160.

WE ARE ALL FAMILIAR WITH THE FACT that laboratory research is making increasing demands on the ingenuity and technical accomplishments of psychologists, among which the need to be familiar with the handling and operating of electronic instruments is paramount. The monograph under review attempts to bring together in one slim volume a brief tour of basic electronic theory, descriptions of the common components used in electronic instruments, descriptions of most of the important types of circuit used, of the instruments in which they are incorporated, and notes on the uses of test instruments, workshop practice, etc. It thus purports to supply sufficient information for the novice to build and operate such instruments for himself. The blurb on the jacket tells us: "No previous knowledge of electronics is required of the reader" and "By presenting electronic theory, circuit design and a variety of instrumentation techniques, this book provides the features of both a text and a handbook." Neither of these claims, however, is tenable without qualification: the reader with no previous knowledge of electronics may find the introduction to the theory manageable, but would undoubtedly find the descriptions of circuitry, etc., very heavy going without additional background reading. Technical words are occasionally introduced without definition, for example, "dielectric" and "degenerative feedback"; these two as a matter of fact are not too important in the contexts in which they appear, but would confuse the novice, as would the wrong numbering of some diagram references. The second claim, that the book is both a text and a handbook, also is too sweeping. In fact, one might say that, in trying to provide the features of both, the book provides neither very adequately. Inadequacies as a text have been mentioned above; as a book of reference the principal shortcoming—almost inevitable in a book of this size—is the failure to specify in sufficient detail the values for components of many of the circuits described. As an example, it would require either a fair knowledge of theory, or a deal of experimentation, to construct a filter circuit with specified characteristics on the basis of the information supplied in the text.

The book has many positive features which deserve mention; it is probably fair to say that, for those with some familiarity with physics and electronics, the book provides a good "orientational" introduction to their applications for particular purposes in psychology, physiology,

and related fields. There are many useful ideas and suggestions for applications which, so far as this reviewer knows, have not been brought together in one book previously. If one wants to know a convenient way of setting up a measuring or recording device the requisite information will probably be found here—in this respect the chapters on stimulus generators and output transducers, and on timing circuits, will be particularly helpful for psychologists—but for adequate details of components, construction, and operation, one will have to enquire elsewhere. However, there is a useful list of places where such enquiries may be made, namely, a list of commercial sources of instruments and components (not exhaustive, of course), a list of the types of tube the authors have found most useful, a list of selected periodicals, and other references. There is an index which is short, but adequate. Two chapters, on test instruments and the laboratory workshop, tell the researcher what he will need to set up his own construction shop. This “do it yourself” section again is useful as an orientational medium, full of useful bits of information about what to buy and where. There is a chapter on transistor theory and applications, a field which will undoubtedly become increasingly prominent in research. Other chapters not previously mentioned discuss different types of tube, power supplies, amplifiers, oscillators, switching circuits, the latter including some discussion of relay applications and relay algebra.

One may conclude that this book would be useful in any laboratory where research requiring instrumentation is carried out, but would require considerable ancillary sources of information where an electronics engineer or technician familiar with psychological research is not on hand.

PETER DODWELL

*Queen's University*

*Studies in Mathematical Learning Theory.* Edited by ROBERT R. BUSH and WILLIAM K. ESTES. Stanford: Stanford University Press. 1959. Pp. viii, 432. \$11.50.

THIS BOOK IS A COLLECTION of papers which were an outgrowth of the psychology of learning workshop held for eight weeks in 1957 at Stanford University as part of a Summer Institute on the Application of Mathematics in Social Science Research. Of the 15 authors of the 20 papers, the major contributors are the editors, Bush and Estes. The other contributors are Anderson, Atkinson, Bower, Burke, Galanter, La Berge, Luce, Mosteller, Papper, Restle, Sternberg, Suppes, and Tatsuoka.

"By *mathematical learning theory* we mean to denote the growing body of research methods and results concerned with the conceptual representation of learning phenomena, the mathematical formulation of assumptions or hypotheses about learning, and the derivation of testable theorems." However, as the editors point out, it is not intended that the reader should regard these several papers together with earlier studies to be found in the literature as constituting a comprehensive general theory of learning to be evaluated in terms of the extent of agreement between the model and some ideal, Utopian theory which would be all-encompassing. Rather, they suggest that the creation of a satisfactory mathematical theory of learning must be the work of "many men over many years."

Work on mathematical learning theory has tended to be along one of two lines; (1) development of a model to represent the stimulus environment and the manner in which changes in this influence behaviour (stimulus-sampling models of Estes *et al.*), and (2) development of a model for the time sequences of responses and for the stochastic processes defined by those models (stochastic learning models of Bush and Mosteller *et al.*). The two lines of endeavour are not in conflict but are essentially equivalent in form and represent attempts to deal with two different aspects of the same problem.

The book is generally organized in terms of these two approaches. Part I is concerned with extensions and adaptations of stimulus-sampling models. Chapter 1 distinguishes between component and pattern models. Chapter 2 deals with stimulus elements not conditioned to any response. Chapter 3 presents an extension of the theory to two-person interaction. Chapters 4, 5, 6, and 7 are applications to serial discrimination, mediated generalization, vicarious trial-and-error behaviour, and latency distributions. Part II (chapters 8-14) is concerned with the stochastic properties of linear models and Part III (chapters 15-20) includes several extensions of the general stochastic model and comparisons of them. The last chapter is a general survey and classification of learning models proposed to date.

Whatever else may be said of the models discussed in this book, they do have a usefulness in subsuming large bodies of data for varying conditions under a few concepts as well as providing a plentiful supply of testable hypotheses in the form of theorems derived from them. The literature attests to this usefulness. Further it has been demonstrated that they have applicability to several different content areas in psychology. Such models also provide precise quantitative statements rather than general forms of functions relating performance to relevant learning variables. Characteristics of the learning process which tend to be over-

looked by other formulations (referred to as "fine-grain" processes) can possibly also be examined.

It is unlikely that the reader would read through this book "at a sitting," but he might use it as a reference for a source of a model in a particular experimental situation. The mathematics presented in the papers are for the most part elementary though somewhat cumbersome. Whether they will become too unwieldy as the models are generalized to fit more complex situations than those hitherto considered remains to be determined.

The editors' title may be somewhat misleading since the major emphasis is not experimental but formal and theoretical. Much of the presentation is devoted to detailed analysis of the logical and mathematical characteristics of models proceeding upon various assumptions about the learning process. Selected experimental data are included to illustrate and compare the application of various types of models and new modifications of the basic models.

A. H. SHEPARD

*University of Toronto*

*The Central Nervous System and Behavior.* Edited by MARY A. B. BRAZIER. The Josiah Macy, Jr. Foundation, 1959. Pp. 450. \$5.25.

THIS BOOK IS A COMPLETE RECORD of the Macy conference held in February, 1958. Three widely differing subjects are dealt with: the history of Russian physiology, the history and present status of conditioning, and current work on the application of neurophysiology and pharmacology to the problems of conditioning.

The first section is profusely illustrated. There are more than a hundred full-page plates; mostly portraits of heavily bearded gentlemen but relieved from time to time by more ornamental subjects such as the seal of the Royal Society of London, Catherines I and II of Russia, the first page of the score of Prince Igor, and a fetching sketch of Sechanov's wife. Pavlov and his work merit 29 pictures, 13 of the master.

The middle section starts with a discussion of the political and financial troubles of Pavlov and Bechterev and modulates through their work on conditional reflexes (as usual on these occasions the question of whether the reflexes are *conditioned* or *conditional* is good for several pages of discussion) to a review of the post-Pavlovian work on conditioning. Sperry and Olds warn that the conditioning situation is more complex than at first appears, or than Pavlov assumed, but on the whole this is a

discouraging section for the reader who has reservations about the universal applicability of the conditioning paradigm to learning.

The remainder of the book is devoted to descriptions of current research on conditioning. Doty describes the effects of using electrical stimulation as a conditioned stimulus; and Olds, his work on the use of electrical and pharmacological stimulation of the rat brain as motivation in the instrumental conditioning situation. Lilly describes similar work with other animals. Morrell's use of the electroencephalographic change as conditioned response is discussed along with some related work by John who recorded electrical activity in various parts of the brain during standard conditioning procedures. Galambos describes the very exciting, though not too well-controlled, experiments demonstrating the effect of attention on the amplitude of evoked potentials in the brain. Some effects of lesions on learning are mentioned by Sperry and Teuber, and Yakovlev outlines a rational, but for those familiar with the conventional terminology, confusing new nomenclature for the nervous system. Most of these studies have by now appeared as journal articles; so there is no need for more detailed summaries here.

The avowed aim of Macy Foundation conferences is to promote communication and stimulate creativity among the participants; the published Transactions are intended to allow a wider audience to share the experience, but on the whole they seem to be regarded as a by-product of secondary importance and no time is wasted on editorial selection. The Director's introductory remarks betray some uneasiness regarding the method of presentation and most readers will, I think, confirm his doubts. In its present "total recall" form the document may be invaluable to future historians of science, and even present-day social psychologists interested in small group interactions, but the student of the C.N.S. and Behavior will probably find it exasperating. At first it is amusing and comforting to see how the experts will misinterpret, or apparently not hear at all, some point in an oral presentation, but the cumulative effect of pointless interruptions, that lead nowhere and could so easily have been edited out, puts a severe strain on one's patience.

Another objection is that the publication of rash statements made on the basis of "preliminary results," though they are the lifeblood of the conference proper, may turn out to be booby-traps in the hands of careless or uncritical readers of the Transactions.

On the positive side there are a few good examples of expert criticism and defence of a piece of research (e.g., Doty), and there are certainly more suggestions for research than would be found in the more formal reports of the experiments. Nevertheless, except for those interested in Russian science in the eighteenth and nineteenth centuries, I would say



that students of neuropsychology would be better off with one of the other symposia on the subject that have appeared recently, and which cover roughly the same ground in a more readable way.

P. M. MILNER

*McGill University*

*American Handbook of Psychiatry*. Edited by SILVANO ARIETI. New York: Basic Books, 1959. 2 vols. Pp. c, 2,098. \$25.00.

THIS HANDBOOK is intended to be a comprehensive textbook of psychiatry. An attempt has been made to cover all the topics which are relevant to psychiatry, to include all the major approaches to the subject, and to represent all points of view. This very wide frame of reference explains the considerable size of the book: 2,098 pages in two volumes. The book is divided into fifteen sections, each representing one part of the field. There are one hundred chapters, representing the work of 111 authors. It is clear then, that the book will suffer from many of the disadvantages implicit in so vast a symposium.

It should be stated at the outset, that, apart from a distinctly psychoanalytical bias, Dr. Arieti's standard of editing has been very high indeed. The field has been carefully divided, and the topics well chosen. Each contributor adheres remarkably carefully to his allotted subject, and with so many authors, such a good result can hardly have occurred spontaneously. Nevertheless, however careful the editing, repetitions occur, the standard of writing and of scholarship is uneven, and perfect coherence has not been attained.

Some of the contributions are very good indeed. Lack of space prohibits even a list of the many excellent ones, but, among them, Norman Cameron's chapter on paranoia and paranoid conditions must rank high as a well-written, succinct account. On the other hand, some contributors are extremely obscure. For example, Edward Stainbrook in his chapter "The Community of the Psychiatric Patient" writes (p. 151): "Hence, the adequate specification of any etiologic matrix of behavioral disease requires the description of antecedent or consequent structure-function deficit in the body, its effect upon or its being affected by the organization, history, and functioning of the personality even as the personality, itself, affects and is affected by the human groups constituting its here-and-now relatedness." Social psychiatry may not be a very advanced field, but there seems no need to disguise this by such verbosity.

The main failing of this book is really inevitable. At first one expects from so comprehensive a textbook a very detailed and exhaustive account

of each topic at an advanced level. However, one is rapidly disillusioned. Just as one cannot expect a very detailed account of manic-depressive psychosis in any encyclopaedia in spite of its size because of the many topics discussed, so in this handbook one cannot expect a very detailed account of any particular topic. For example, only 19 pages are devoted to the history and symptomatology of hysteria, and while D. W. Abse's account of this illness is a good one, it is necessarily brief. On the average, each of the 100 chapters is twenty pages in length, including the bibliography, and, as many books are available on virtually each topic for which there is a chapter, each chapter represents a short summary of the work in its particular field. This is not the fault of the present work, but merely a necessary consequence of the breadth covered.

This handbook, then, cannot be regarded as a detailed source book. It must be used more as an encyclopaedia. The term "handbook" here is used more in the sense of a *vade mecum* for the student than in the sense of a definitive work of reference for the scholar.

R. W. PAYNE

Queen's University

*Psychology of Perception.* By WILLIAM N. DEMBER. New York: Henry Holt, 1960. Pp. xi, 402. \$6.50.

IT IS QUITE PROBABLE that if you teach an undergraduate course in perception you do not use a textbook in the ordinary sense of the word for the simple reason that you have not found a suitable one. This may sound paradoxical in view of the plethora of books that psychologists write and perhaps read, and in view of the importance of problems classified as perceptual, but it seems to be true nevertheless.

Dember's book should go a long way towards filling the apparent vacuum. It has grown out of the author's experience in teaching a course in perception and thought—although it contains nothing on the latter—and it covers the material that he himself would like to see included in a one-semester laboratory course in perception. It is hardly likely that everyone would agree with Dember's selection of, and emphasis on, various parts of the subject matter, yet there should be enough concurrence to render the book useful in many classrooms.

The avowed purpose of the book is to present a representative sample of problems, methods, and findings in perception. Perception is defined in terms that are obviously influenced by information theory. Threshold, for instance, is described as "the minimal amount of information required for the accomplishment of a perceptual task" (p. 14). At the same time, however, information theory approach to perception is not followed or

even mentioned. The material covered is strongly anchored to experimental evidence, with a good deal of emphasis on method and with only occasional excursions into the realm of theory. The author has adopted a constructively critical attitude throughout the book achieving a healthy balance between the accomplishments and uncertainties that characterize the whole complex field of perception.

Dember's perceptual world in the book is silent, almost colourless, and detached from the nervous system. Auditory perception does not enter the picture, colour is given one page on colour contrast, and as to the physics and physiology of perception the reader is sent to other sources, such as Geldard. Within these limitations the book lives up to its title quite adequately. The ten chapters deal with topics that range all the way from threshold measurement techniques and visual psychophysics through organization of perception, stimulus context, set, and the role of learning, to problems dear to the heart of the New Look adherent. The final chapter on stimulus change, curiosity, and exploratory behaviour reflects the author's personal interest in this increasingly fashionable area. Although the appropriateness of its inclusion is cogently rationalized it might have been preferable to have had a summary chapter in its stead.

The book easily earns top marks on two accounts, its organization, and its up-to-date references. Both individual chapters and the whole book are thoughtfully composed, without being slanted towards any particular theoretical school of thought. The organization is based on systematic consideration of the selection and the method of presentation of stimuli, and on the subject's task in perceptual experiments. This results in an orderly progression from subthreshold to suprathreshold phenomena and from the effects on perception of relatively simple, individual stimulus variables, to more complex interactions between stimulus variables and other conditions. The terms "sensation" and "perception" may be used here if one believes in the distinction; Dember does not and neither does this reviewer, at least at the time of this writing. Artifices such as the systematic two-by-two categorization of psychophysical methods and the explicit distinction between the experimental tasks of detection, discrimination, and identification, between intramodal and intermodal context effects, and between explicit and implicit set, to mention some examples, not only provide a meaningful frame of reference for the data, but also serve a useful pedagogical function. Concise summaries of individual chapters augment these organizing devices.

The second outstanding strength of the book lies in the fact that approximately one-half of some 200 items referred to in the text date from the last six years or so. This proportion compares very favourably with,

say, Bartley's *Principles of Perception*, published two years ago, where less than one-fifth of all references fall in a similar category. A new record must also have been established by Dember for including in his book that came out in early 1960 ten references to work published in 1959. All this has the effect of bringing to the reader's attention much of the interesting and important work done by various investigators in recent years. Examples are Coombs's methods of scaling similarity, Stevens's work on scaling attributes of stimuli (the most recent work on cross-modality matching, if included, may have provided a needed antidote to Dember's ambivalent attitude towards Stevens's methods), Cohen's Ganzfeld phenomena, the stopped image of Riggs and his associates (the work of Ditchburn and his collaborators is not mentioned), investigations of perception of slant by Clark and his group, McGill studies on sensory deprivation, and the response probability interpretation of tachistoscopic experiments. Incidentally, the authors most frequently mentioned are Dember, Postman, Berlyne, and Bruner.

Dember's style of writing is clear and his language should not present difficulties to typical undergraduates. Instructors who contemplate using this text will be glad to know that there are enough sections in the book that require clarification or knowledge of related matters. In this sense the book is no exception but then there never have been nor ever will be any perfect textbooks. The physical features of the book are quite attractive as long as you do not remove the dust jacket: the silvery covers are eerie and the colour rubs off on one's fingers. This, however, should not deter you from looking inside the book. Whether you are searching for a good text or simply wish to brush up on perception this book may well be the best one to begin with.

ENDEL TULVING

*University of Toronto*

